

LA-UR-18-25316

Approved for public release; distribution is unlimited.

Title: Combinatorial Optimization on D-Wave

Author(s): Coffrin, Carleton James

Intended for: Web

Issued: 2018-06-18

Disclaimer:

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the Los Alamos National Security, LLC for the National Nuclear Security Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396. By approving this article, the publisher recognizes that the U.S. Government retains nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.



advanced network
science initiative
(ansi)

Combinatorial Optimization on D-Wave

Carleton Coffrin

A-1

UNCLASSIFIED

What is Computational Optimization?

What is Optimization?

A Mathematical Program

$$\min : f(x)$$

Objective Function

What is Optimization?

$$\min : -3x + y + z \Rightarrow -0.5$$

s.t.

$$x^2 \leq yz$$

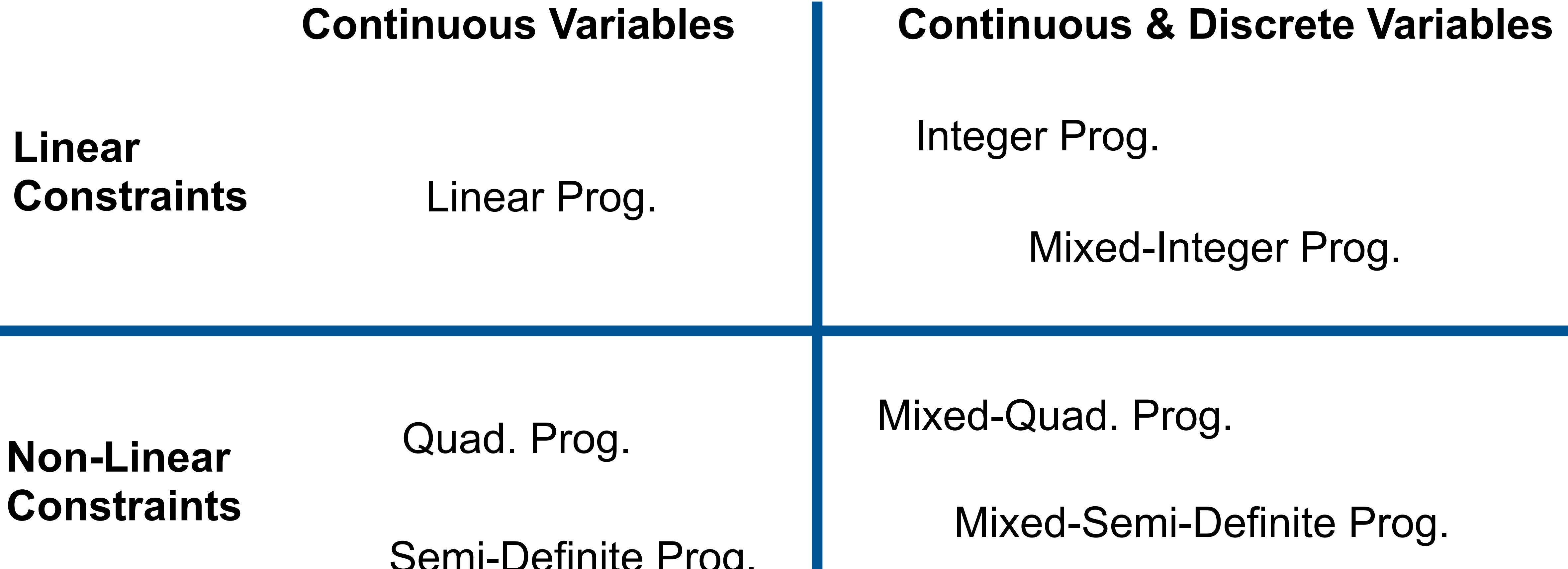
$$x^2 + y^2 \leq -z + 1$$

$$y \geq 0, z \geq 0$$

$$x \in \mathbb{R}$$

$$x, y, z = \frac{1}{2}$$

Classes of Optimization Problems



What is Combinatorial Optimization?

What is Combinatorial Optimization?

Good News

assignment space is finite

$$\min : f(x)$$

Combinatorial Object

Objective Function

$$x \in \{0, 1\}^n$$

Binary States

$$2^n$$

Bad News

assignment space is HUGE!

$$x \in \mathbb{Z}^n$$

Integer States

$$d^n$$

$$x \in \mathcal{G}_n$$

Graphs with n nodes

$$2^{\binom{n}{2}}$$

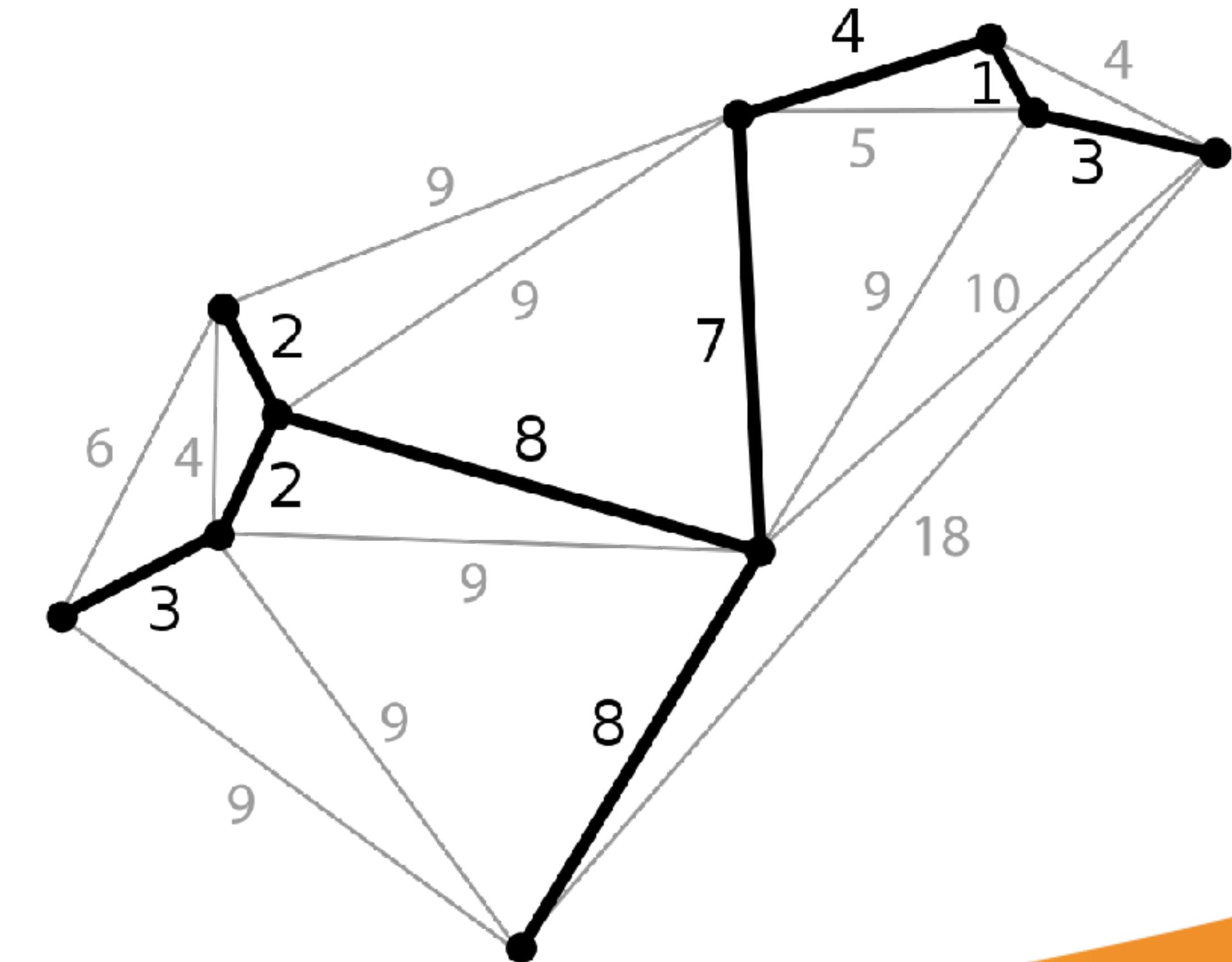
Why Combinatorial Optimization?

- All over the place
- Beautiful Mathematics and Structure
- Core building blocks in more complex optimization algorithms
- Many are hard to solve (NP-Hard)
 - BUT, not necessarily hard

Minimum Spanning Tree

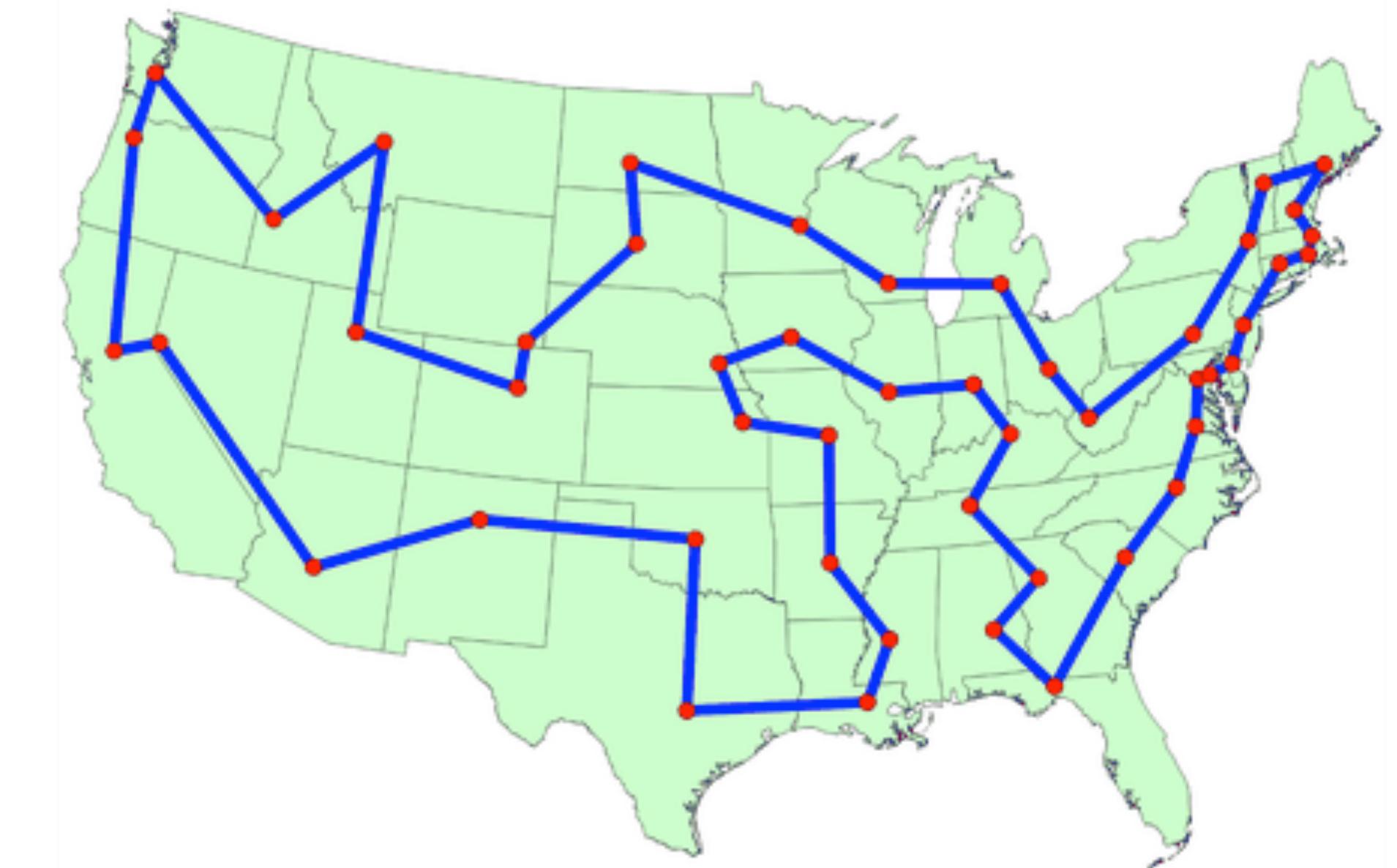
- Given a weighted graph
- Find a subtree with minimum edge weights

$$O(|E| \log(|E|))$$



Traveling Salesman

- Visit all points, minimize distance
- $f(x)$ - euclidian distance
- x - set of all tours
 - (hamiltonian cycles)



Knapsack

- Take items to maximize profit
- Limited capacity you can carry (c)

$$\min : \sum_i p_i x_i$$

s.t.

$$\sum_i w_i x_i \leq c$$

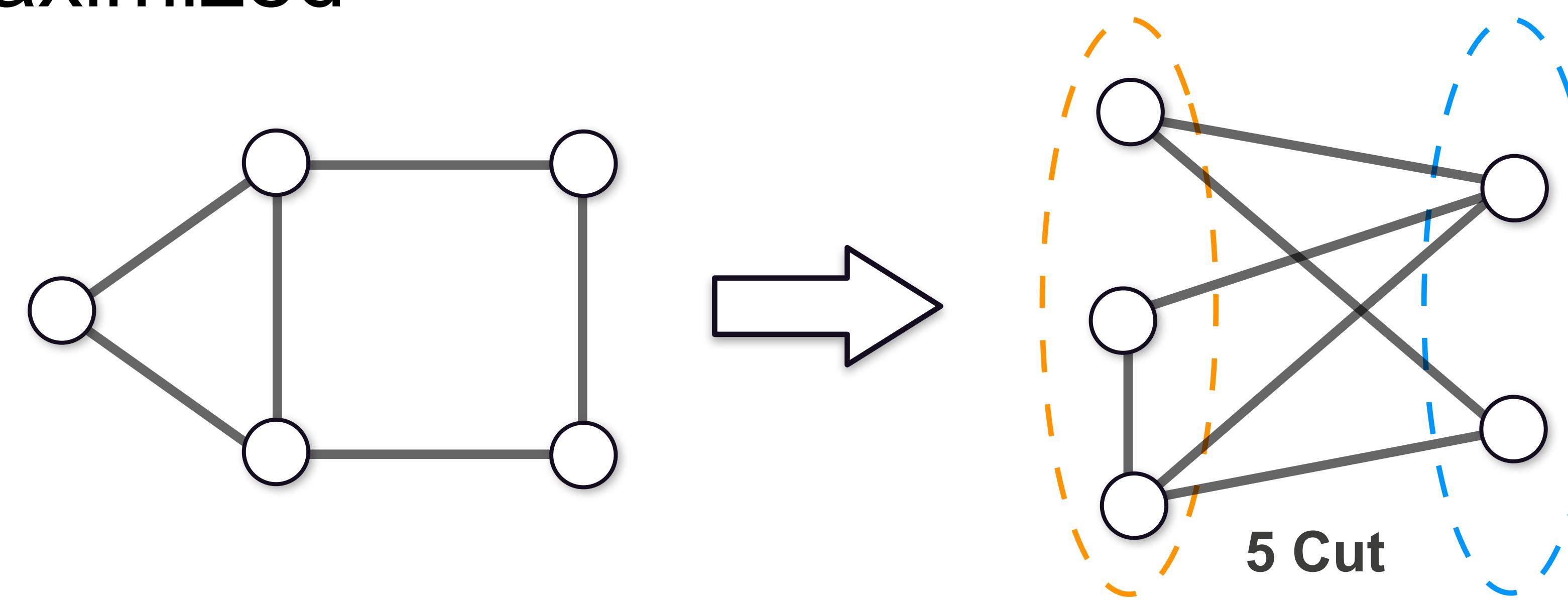
Combinatorial
Challenge



$$x \in \{0, 1\}^n$$

Max-Cut

- Given a graph
- Partition into two sets, such that edges between is maximized



Why is Combinatorial Optimization Scientifically Interesting?

Real World Motivation

- Many practical applications
 - 1D, 2D, 3D bin packing
 - routing
 - scheduling
 - resource allocation

Comb. Opt. needs to be solved



You must have a solution, but...
good algorithms are not obvious
(P vs NP)

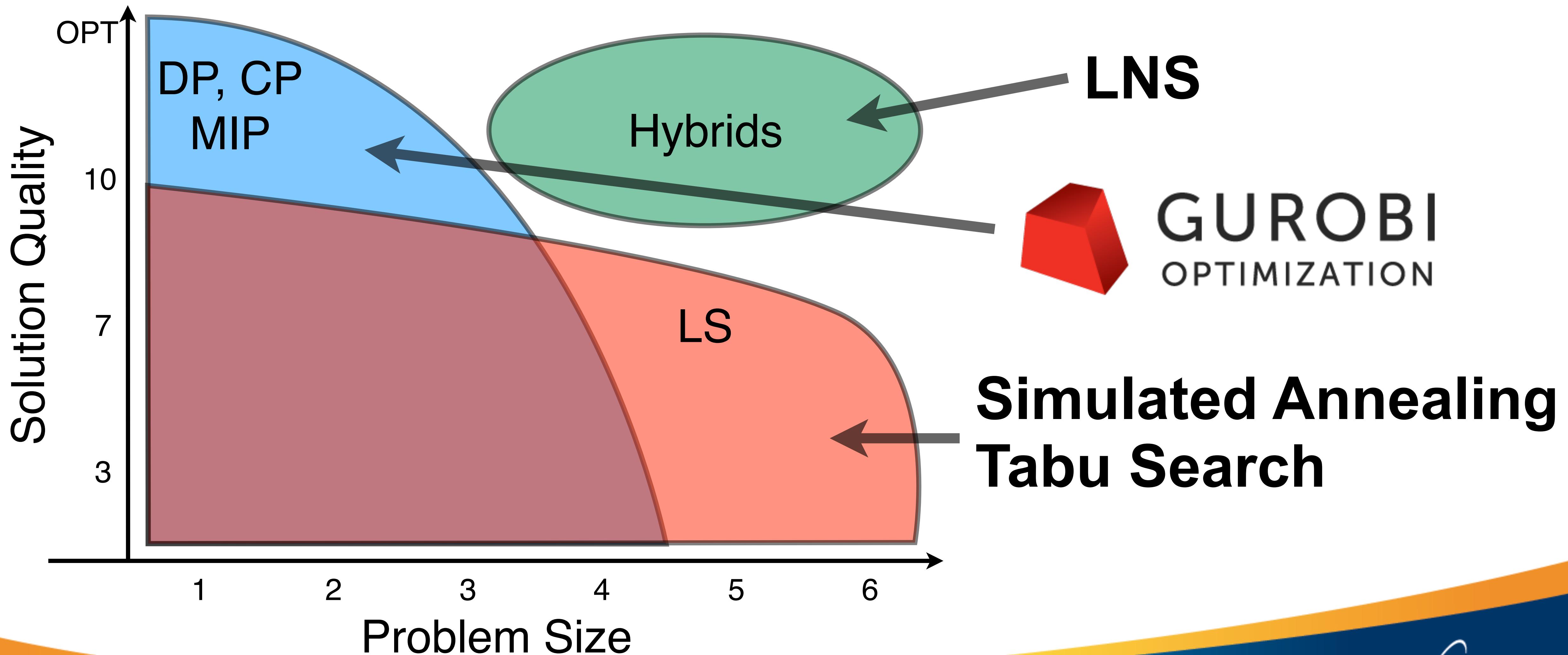
Assumptions of the Practitioner

- $P \neq NP$ (stuck with exp. scaling)
- Real-world data is not adversarial
 - i.e. worst-case complexity is too rough a measure
- Problem structure can be exploited on a case-by-case basis
- Even the structure of specific input datasets

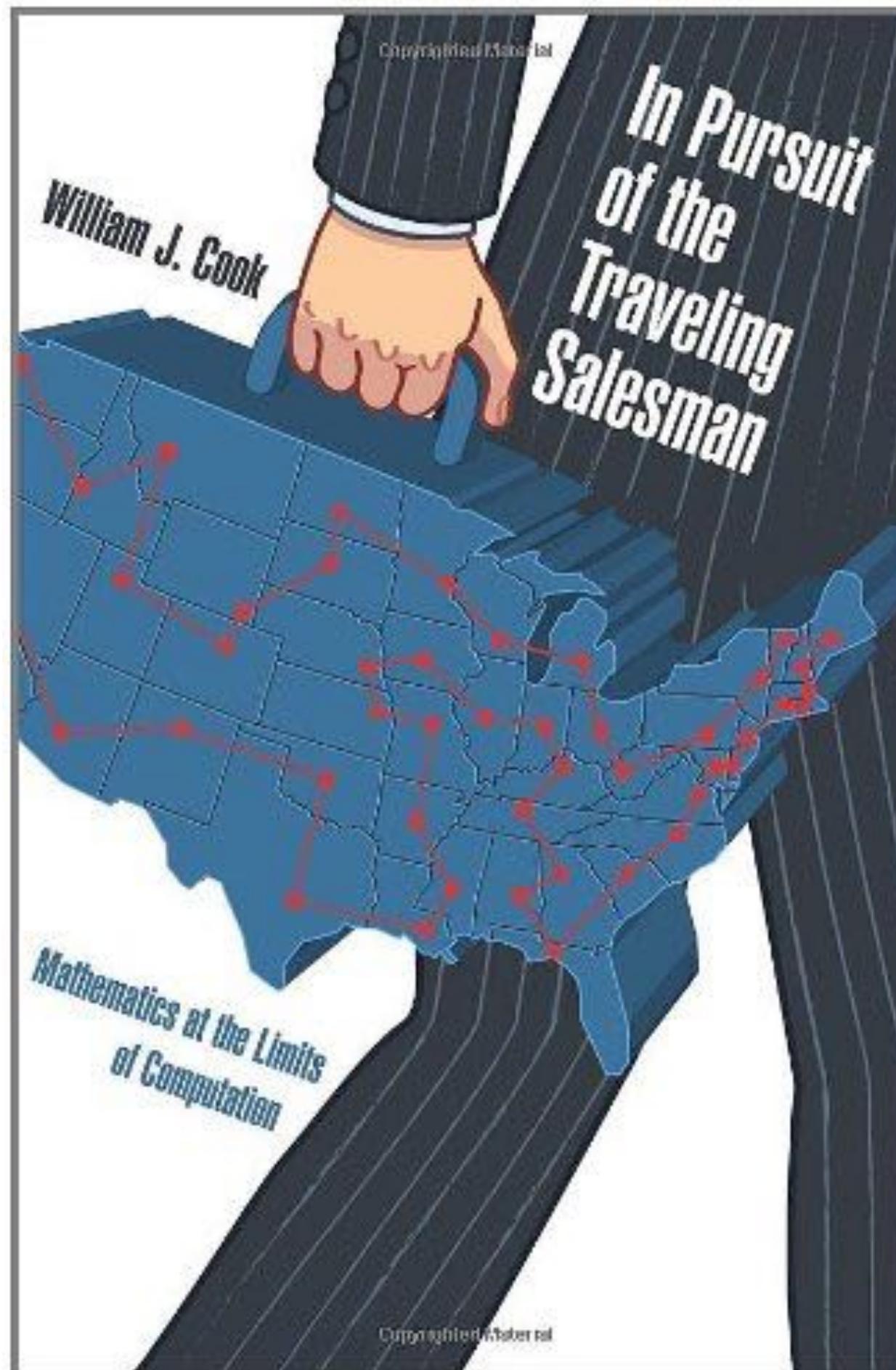
Solution Methods

- Local Search (LS) (a.k.a. heuristics)
 - Simulated Annealing, Tabu Search, ...
- Integer Programming (IP)
 - Applied Math (simplex alg.)
- Constraint Programming (CP)
 - logical deduction, dynamic prog.

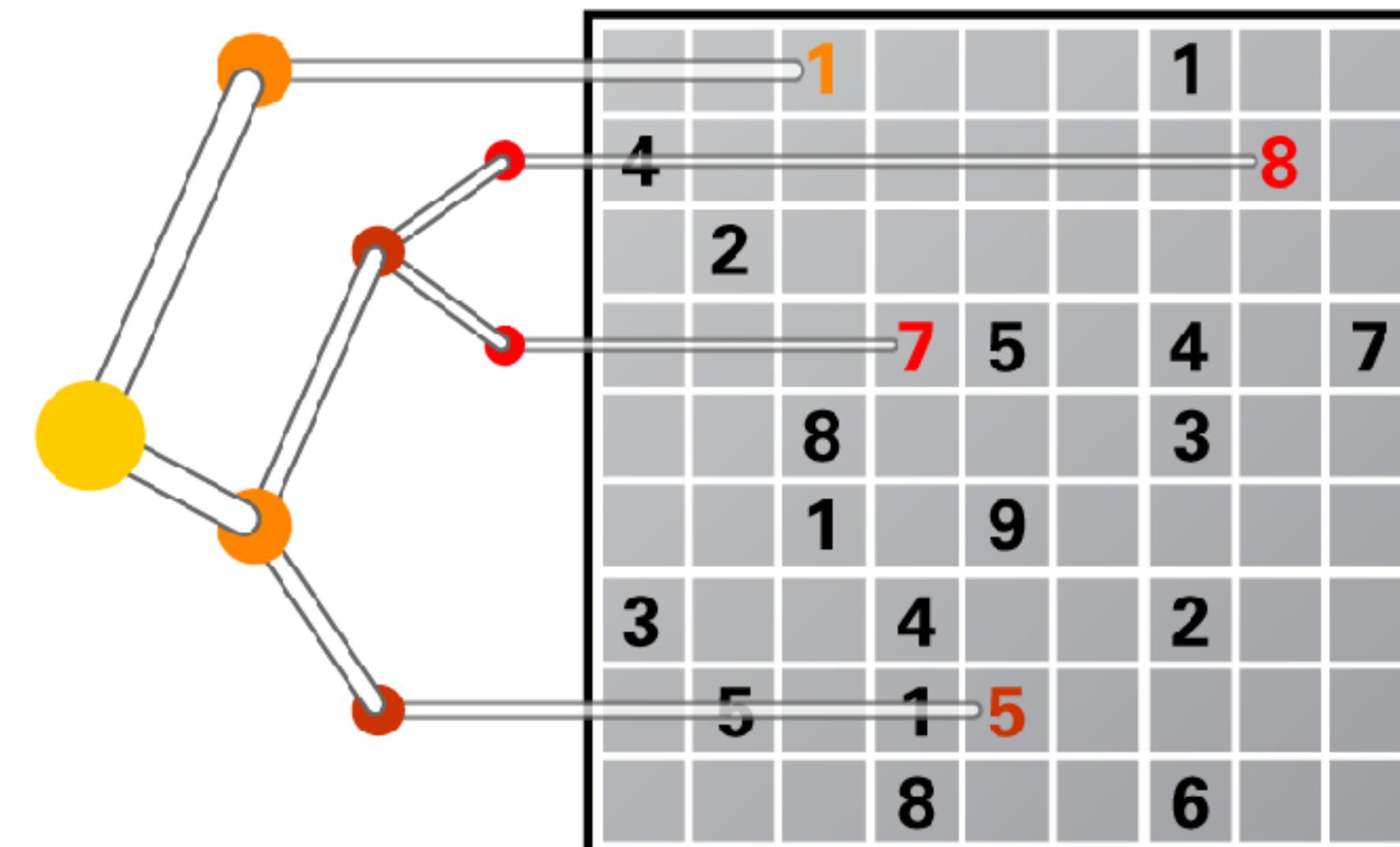
Solution Method Classes



Homework

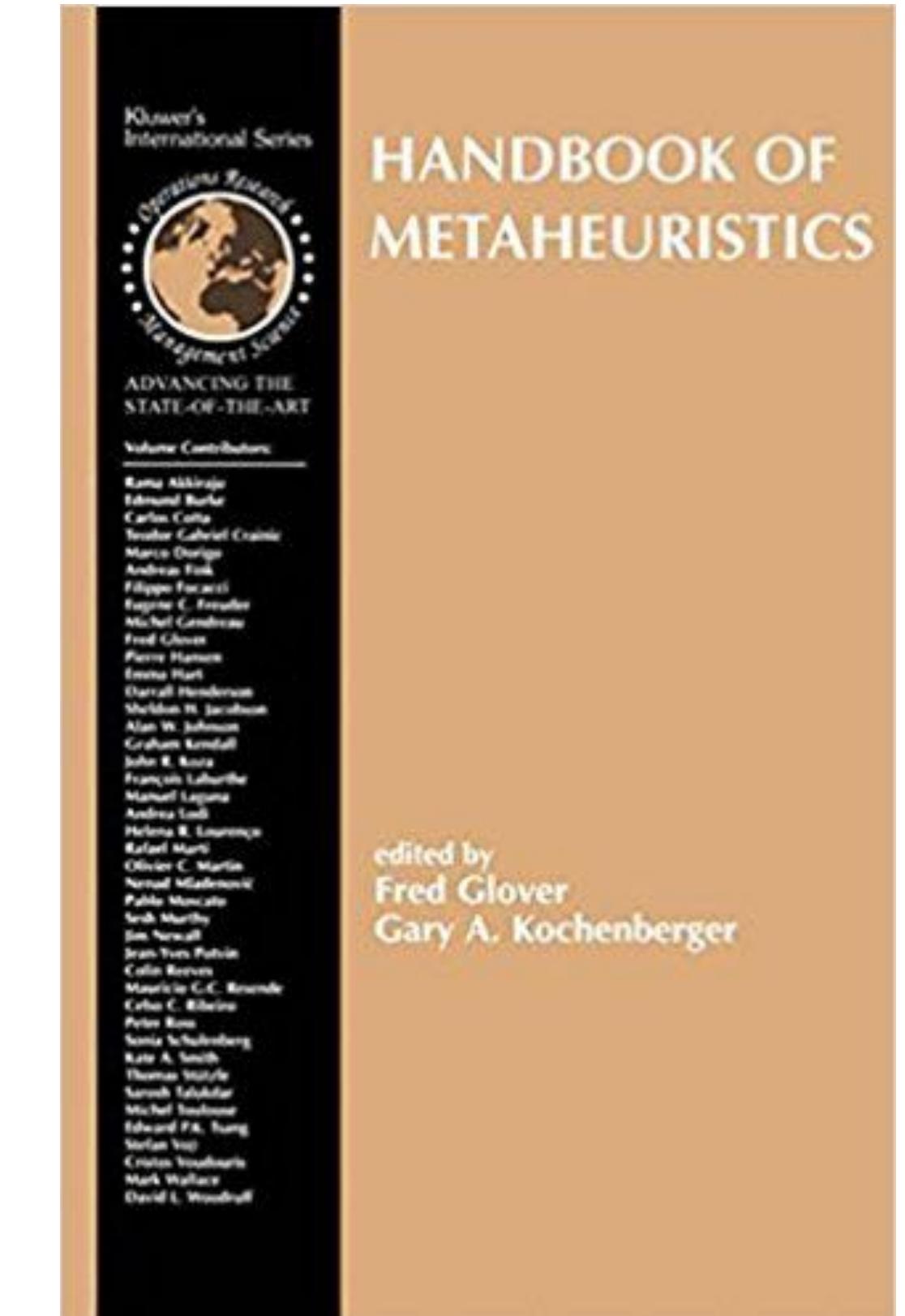


coursera



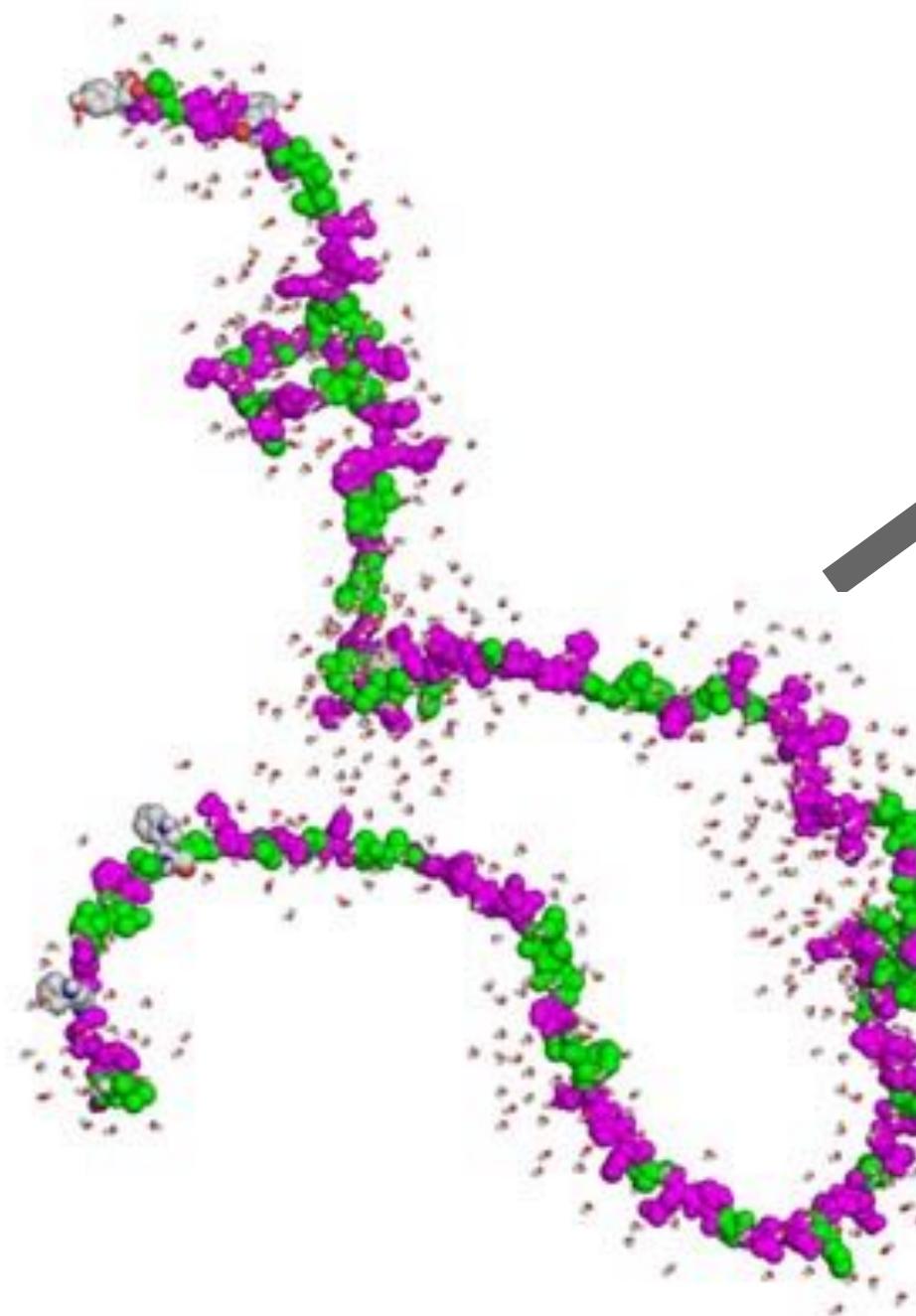
Discrete

Simulated Annealing (1 of 19)



Algorithm R&D for NP-Hard Problems?

Algorithm Benchmarking



**Protein
Folding**

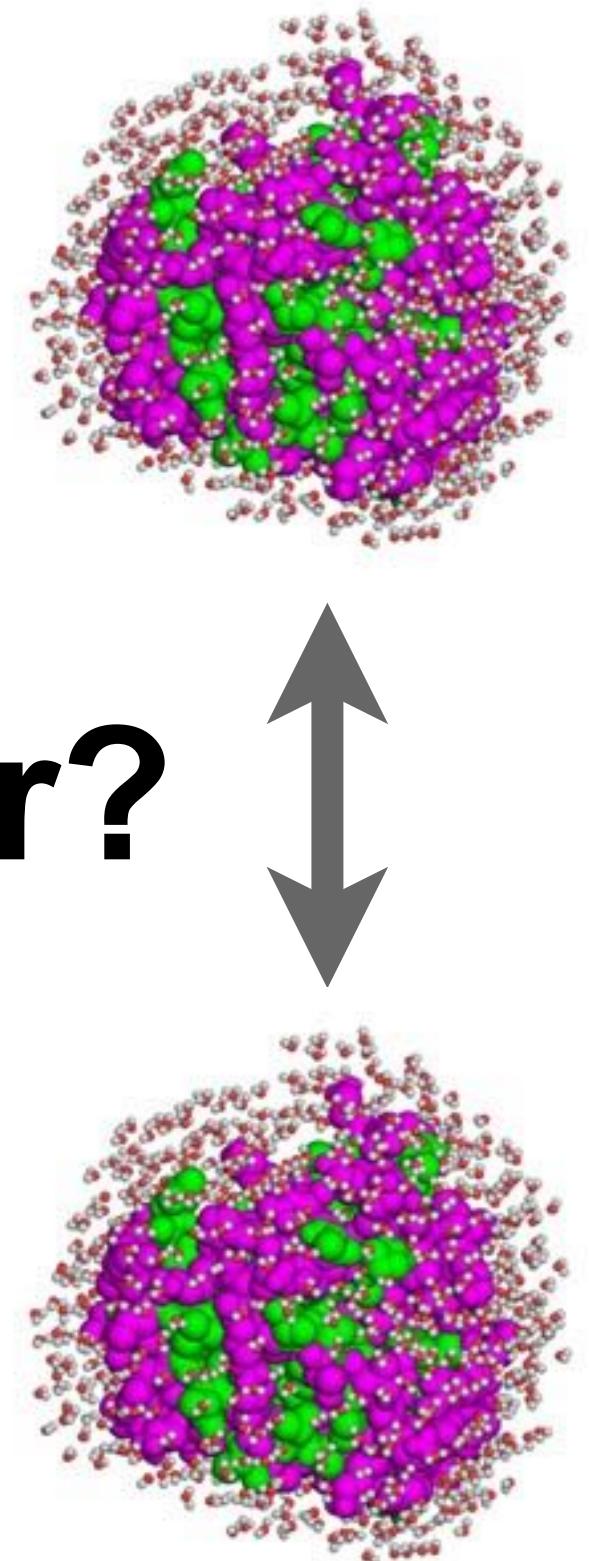
Current State-of-the-Art

Claim: On this problem,
The proposed algorithm is state-of-the-art.
The ends justify the means.

Proposed Algorithm

Not Claiming: On this problem,
the proposed algorithm is the best
(or even asymptotically better)

Better?



Experimental Algorithm Design

- Representative collection of specific test cases
 - Usually from industry applications
 - Test multiple algorithms on such cases
 - Compare key metrics
 - e.g. solution quality, wall-clock time

So what! The algorithm is still exponential...

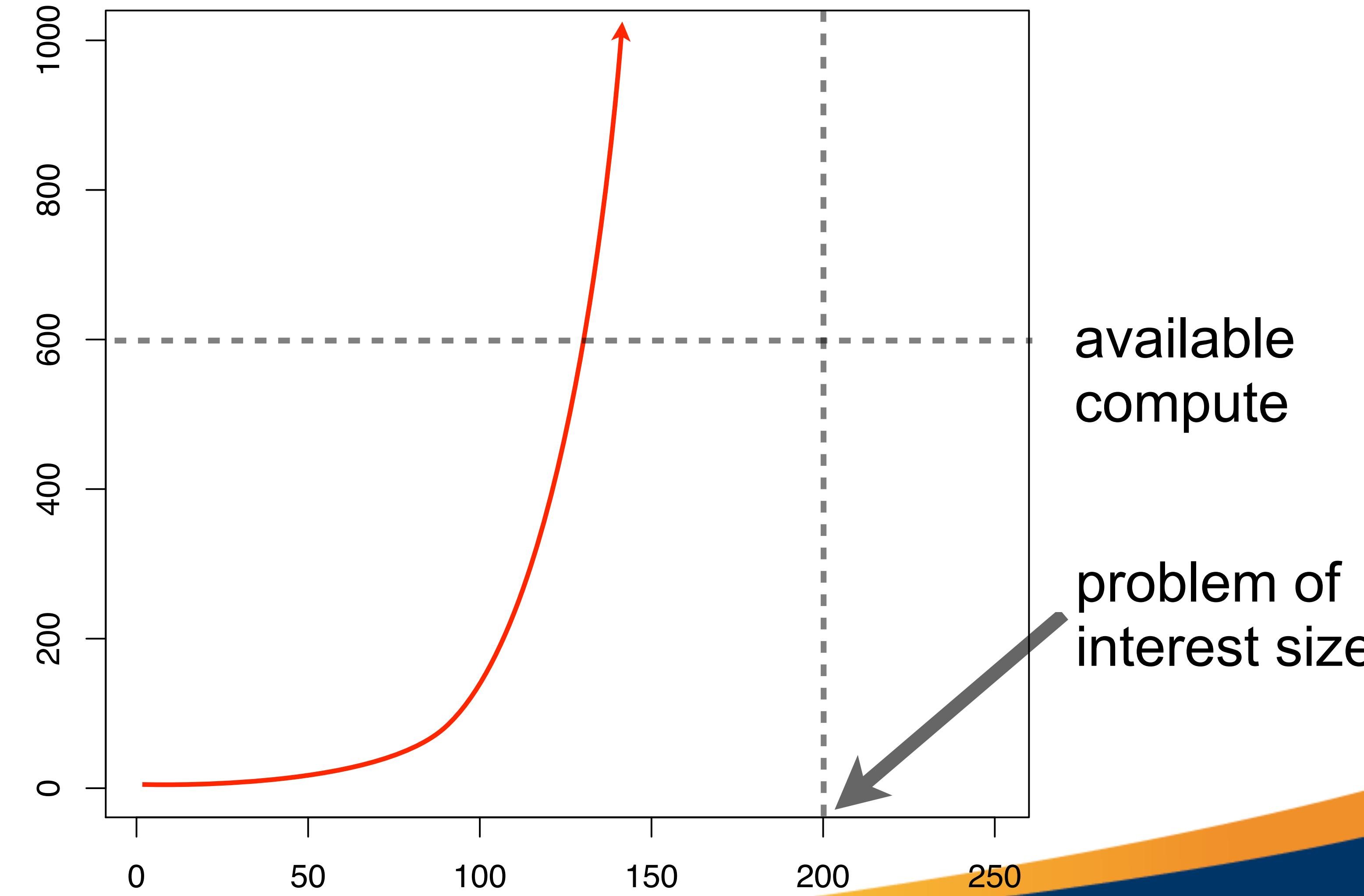
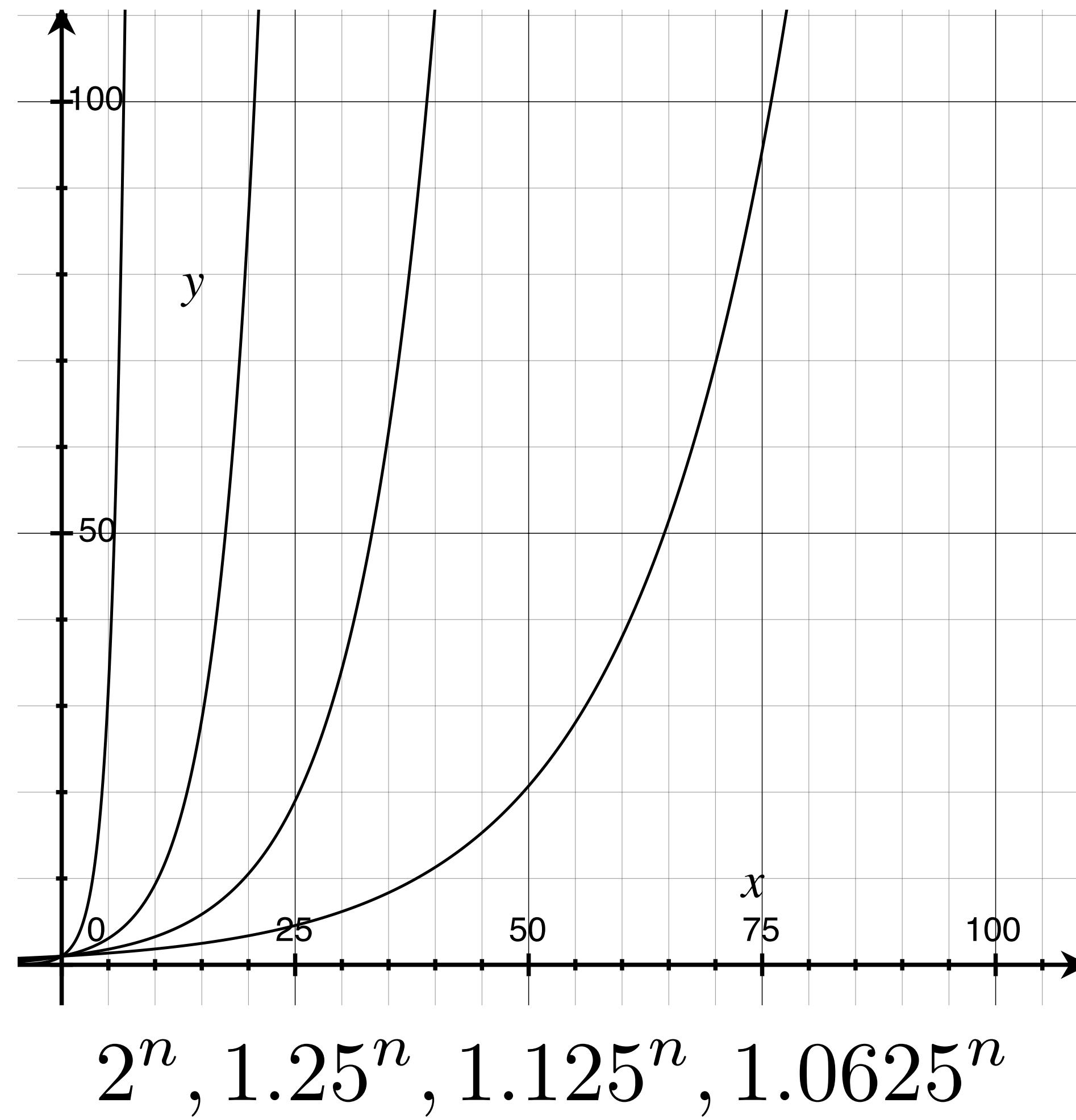
Compute budget: 10^{10}

$$10^n \Rightarrow n \leq 10$$

$$10^n \Rightarrow 2^n$$

$$2^n \Rightarrow n \leq 33$$

Embrace Exponential



Real Progress

- Mixed-Integer Programming
 - Includes Combinatorial Opt.



**500,000x *algorithmic*
speedup from 1993-2013**

<http://bob4er.blogspot.com/2015/05/amazing-solver-speedups.html>

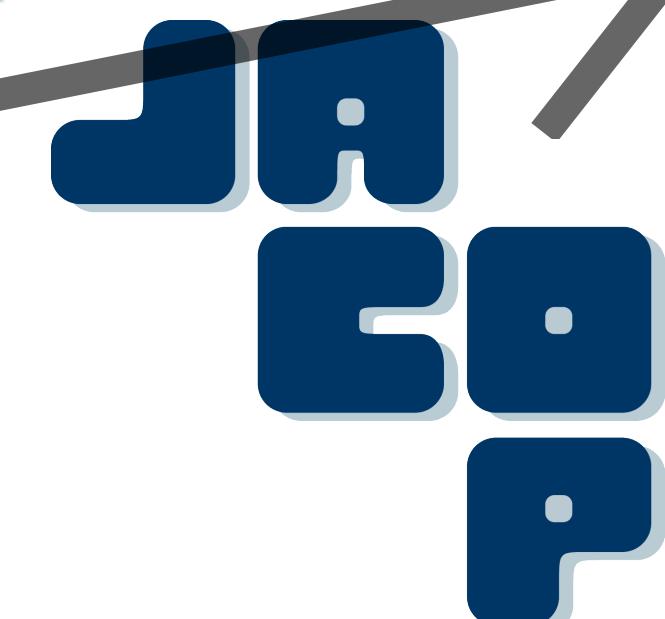
300,000x computational

The Optimization Landscape

Solvers (i.e. Algorithms)



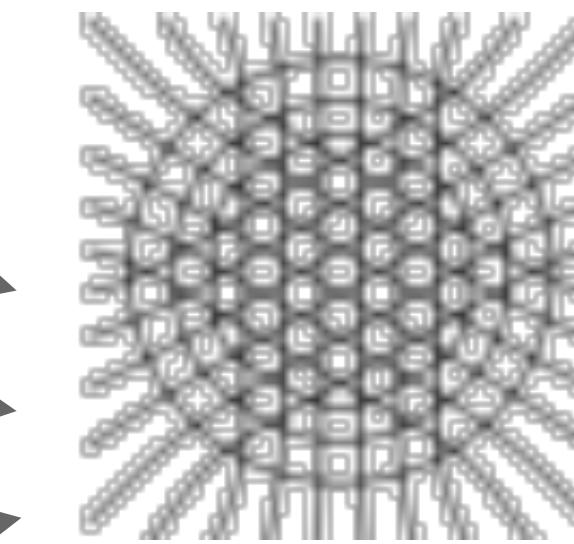
CPLEX



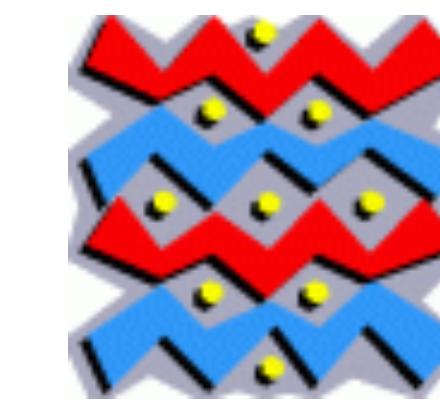
Benchmarks (i.e. Problems)



CSPLIB



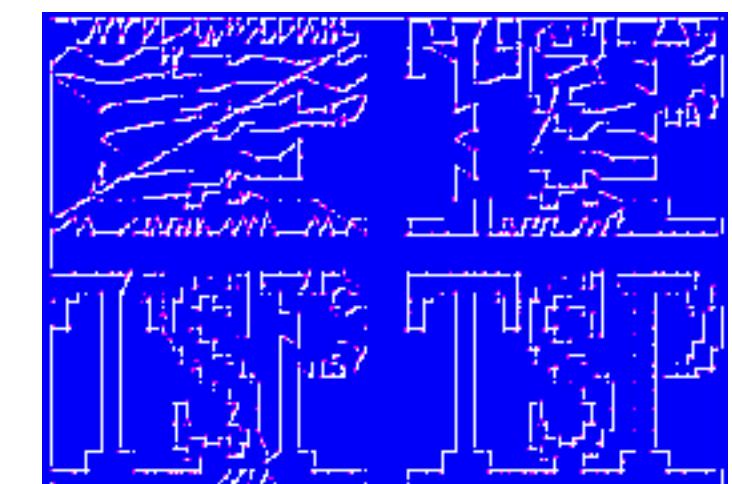
MIPLIB



MINLPLIB

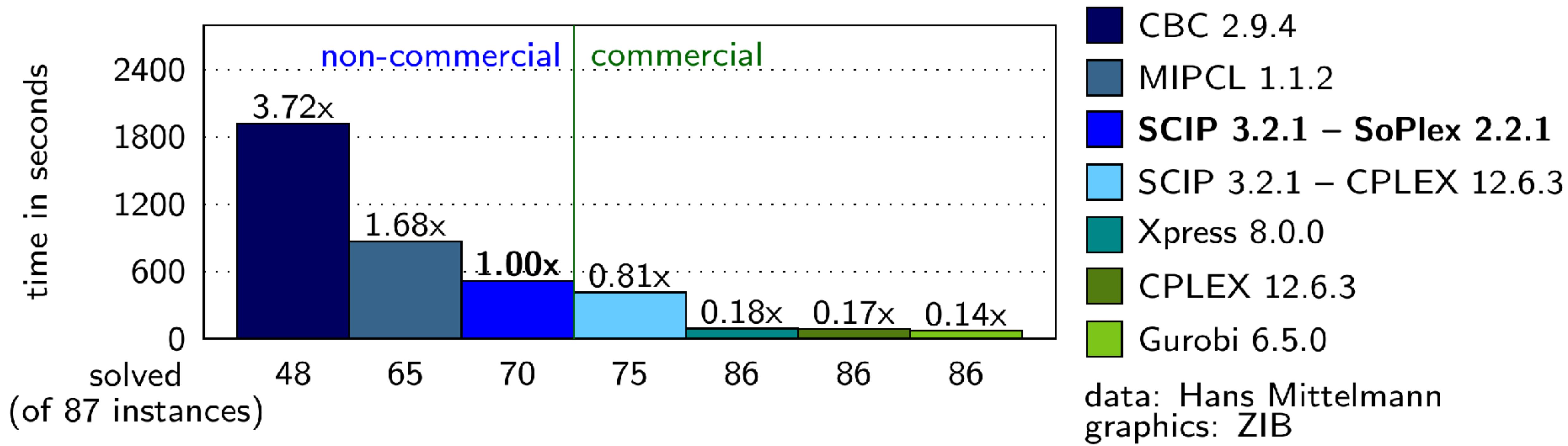


QPLIB



TSPLIB

Optimization Science



<http://plato.asu.edu/ftp/milpc.html>

<http://scip.zib.de/>

So what about D-Wave? (QA)

First-Order Approximation

- Input a classical hamiltonian (J, h)
- If we set a sufficiently long annealing time
- We get the ground state of an Ising Model

$$H(\sigma) = \sum_{i,j} J_{ij} \sigma_i \sigma_j + \sum_i h_i \sigma_i$$
$$\sigma \in \{-1, 1\}$$

An Optimization View

- Binary-Quadratic Program
- No constraints (QUBO)

Eliminates a number
of interesting problems
(e.g. Knapsack, TSP)

Class of Combinatorial Optimization

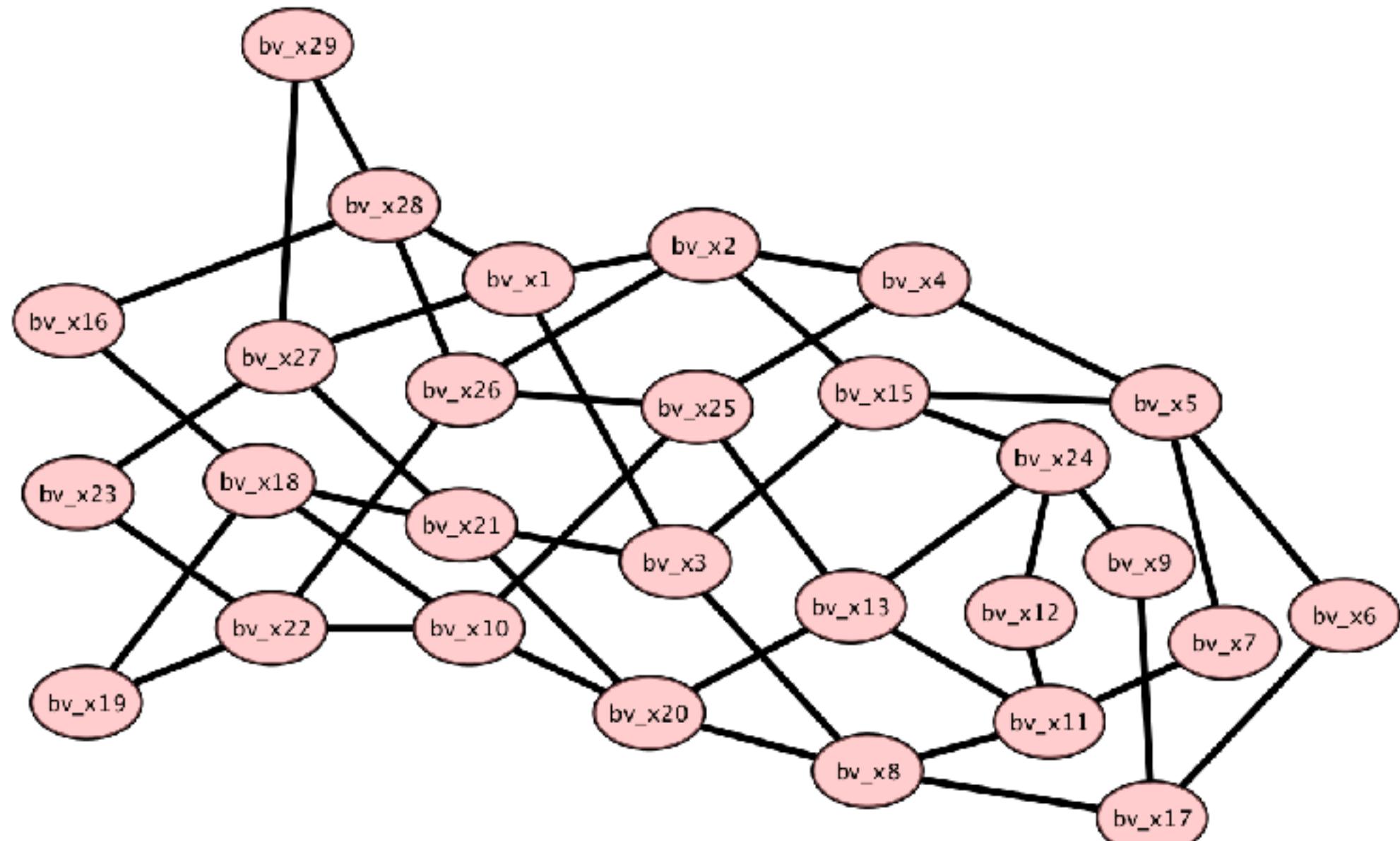
$$\min : \sum_{i,j} J_{ij} \sigma_i \sigma_j + \sum_i h_i \sigma_i$$

s.t.

$$\sigma \in \{-1, 1\}$$

What can we solve?

An Optimization View



$$\mathcal{G} = (\mathcal{N}, \mathcal{E})$$

MAX CUT = Anti-Ferromagnet

$$\min : \sum_{i,j \in \mathcal{E}} \sigma_i \sigma_j$$

s.t.

$$\sigma_i \in \{-1, 1\} \quad \forall i \in \mathcal{N}$$

Ising-Friendly Comb. Opt. Problems

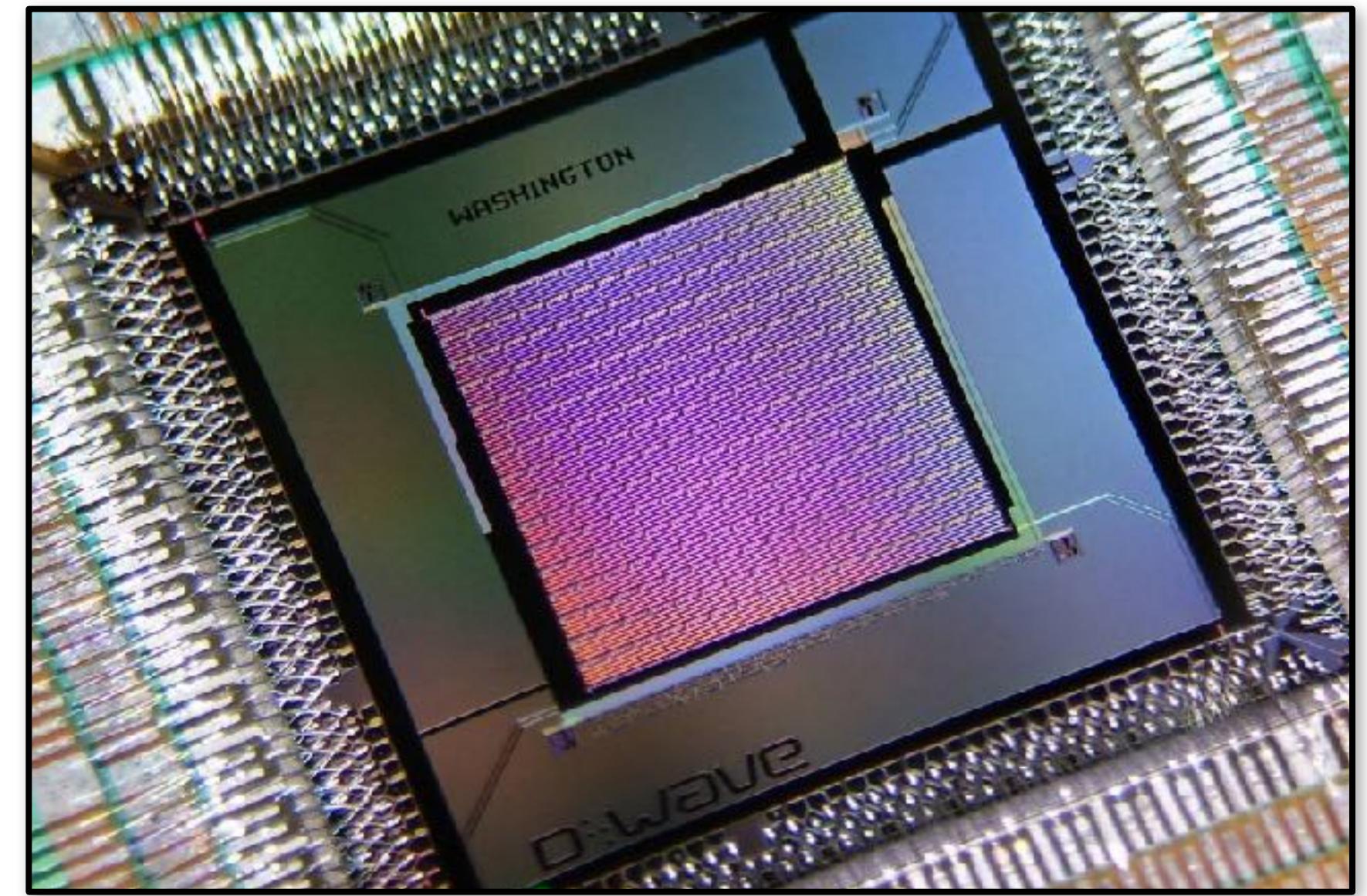
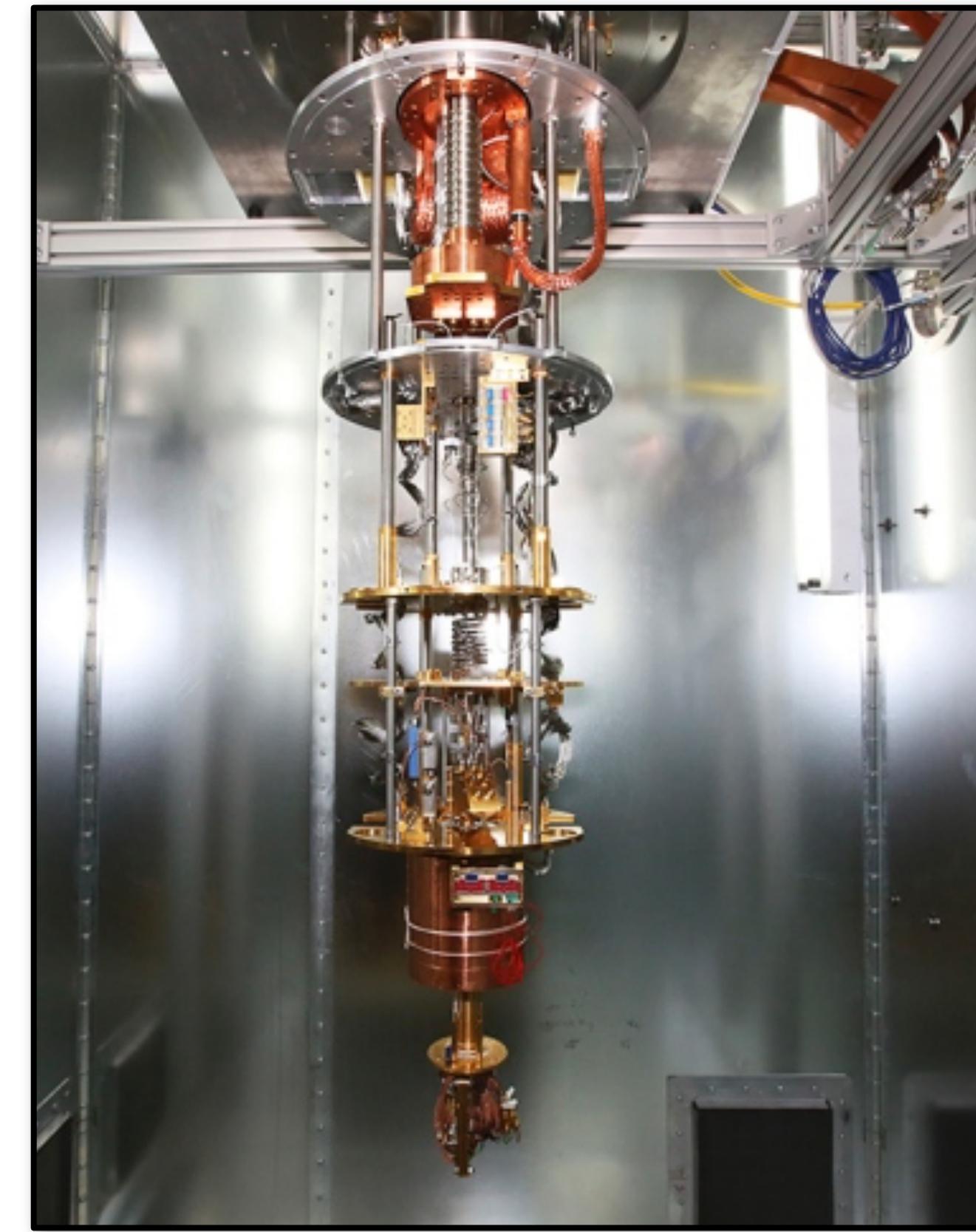
- Weighted Max-Cut
- Max Clique
- Max Independent Set
- Vertex Cover
- ...

$$\begin{aligned} \min : & \sum_{i,j \in \mathcal{E}} J_{ij} \sigma_i \sigma_j + \sum_{i \in \mathcal{N}} h_i \sigma_i \\ \text{s.t.} : & \sigma_i \in \{-1, 1\} \quad \forall i \in \mathcal{N} \end{aligned}$$

Ising formulations of many NP problems

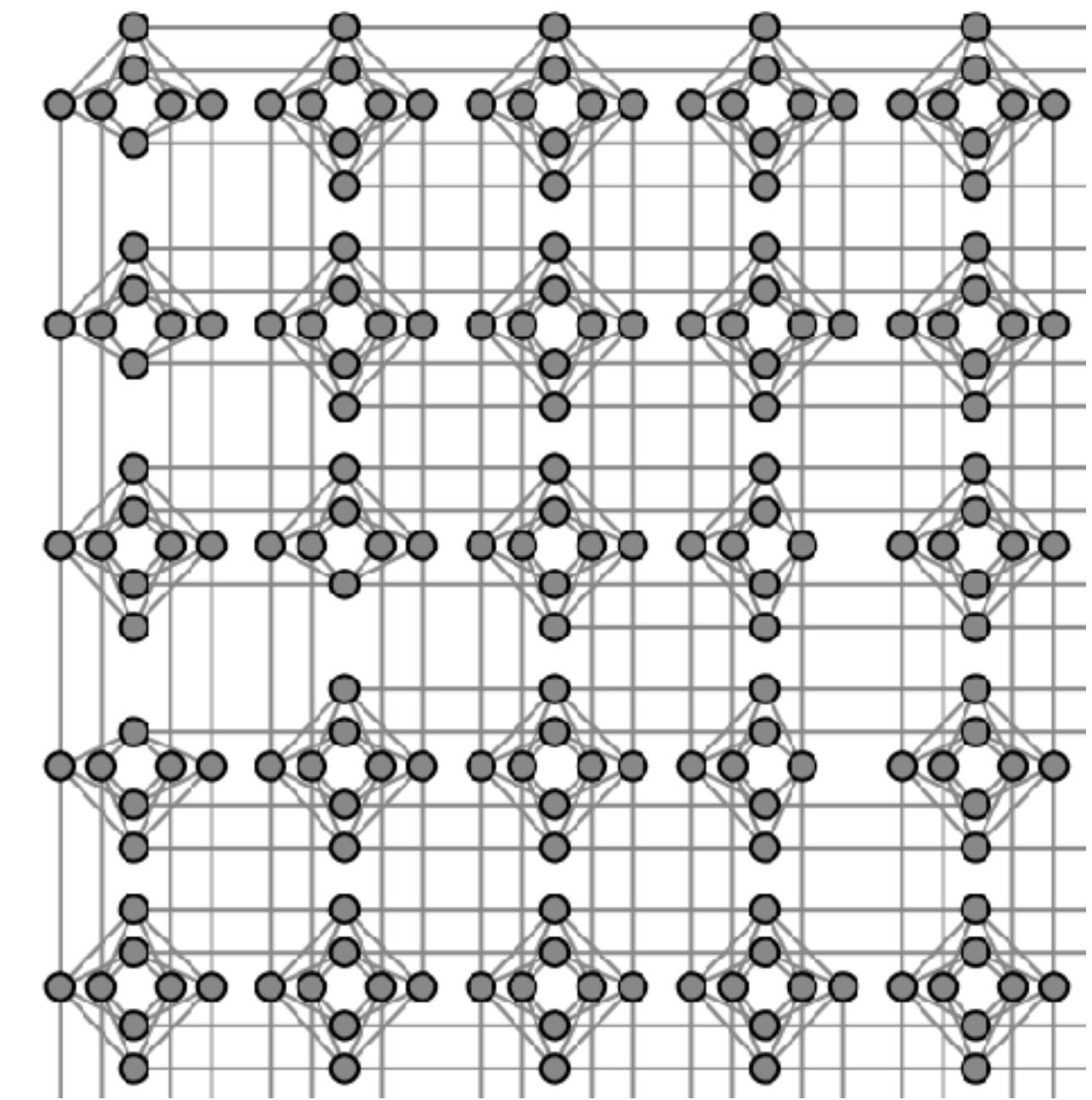
<https://arxiv.org/abs/1302.5843>

This thing is amazingly complex



Challenges

- Hardware Connectivity Graph (Chimera)
- Sparse
- Bi-partide
- max-cut is trivial

 C_5

Challenges

- The D-Wave Problem

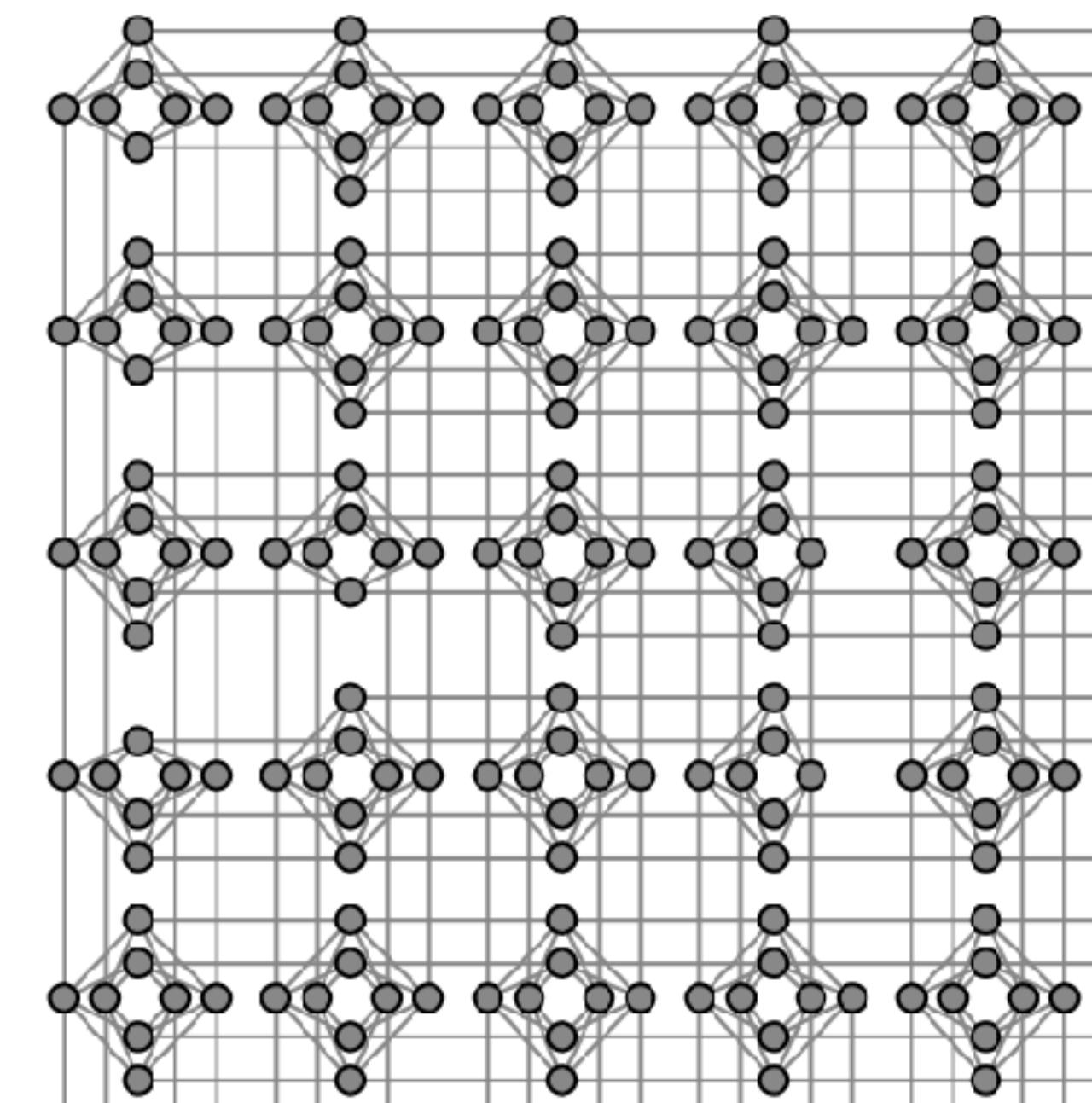
$$\min : \sum_{i,j \in \mathcal{E}} J_{ij} \sigma_i \sigma_j + \sum_{i \in \mathcal{N}} h_i \sigma_i$$

s.t.

$$\mathcal{E} \subseteq \mathcal{C}_{16}$$

$$\begin{aligned} -1 \leq J_{ij} &\leq 1 & \sigma_i \in \{-1, 1\} \quad \forall i \in \mathcal{N} \\ -1 \leq h_i &\leq 1 \end{aligned}$$

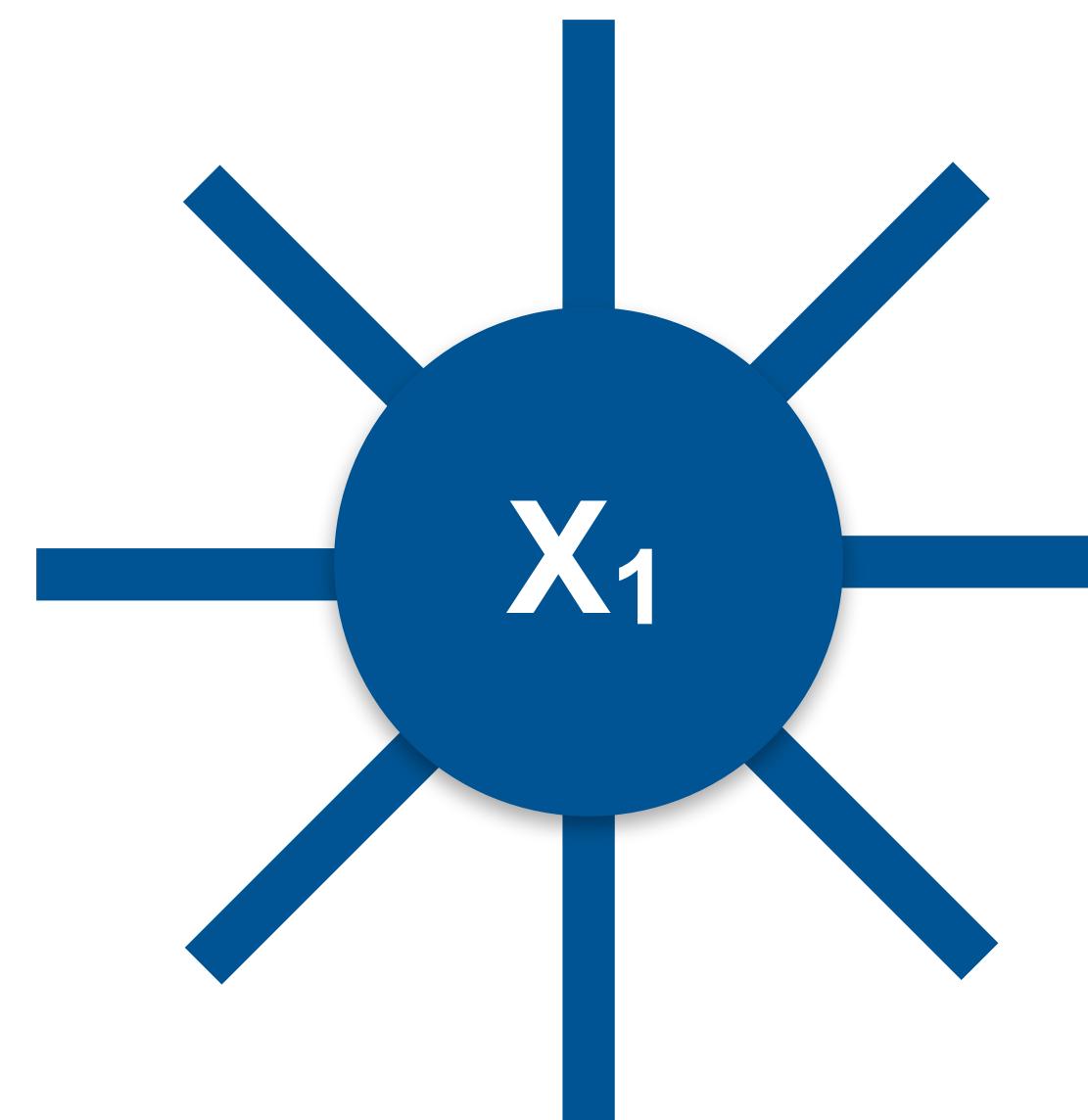
$$\begin{aligned} \min : & \sum_{i,j \in \mathcal{E}} J_{ij} \sigma_i \sigma_j + \sum_{i \in \mathcal{N}} h_i \sigma_i \\ \text{s.t.} : & \sigma_i \in \{-1, 1\} \quad \forall i \in \mathcal{N} \end{aligned}$$



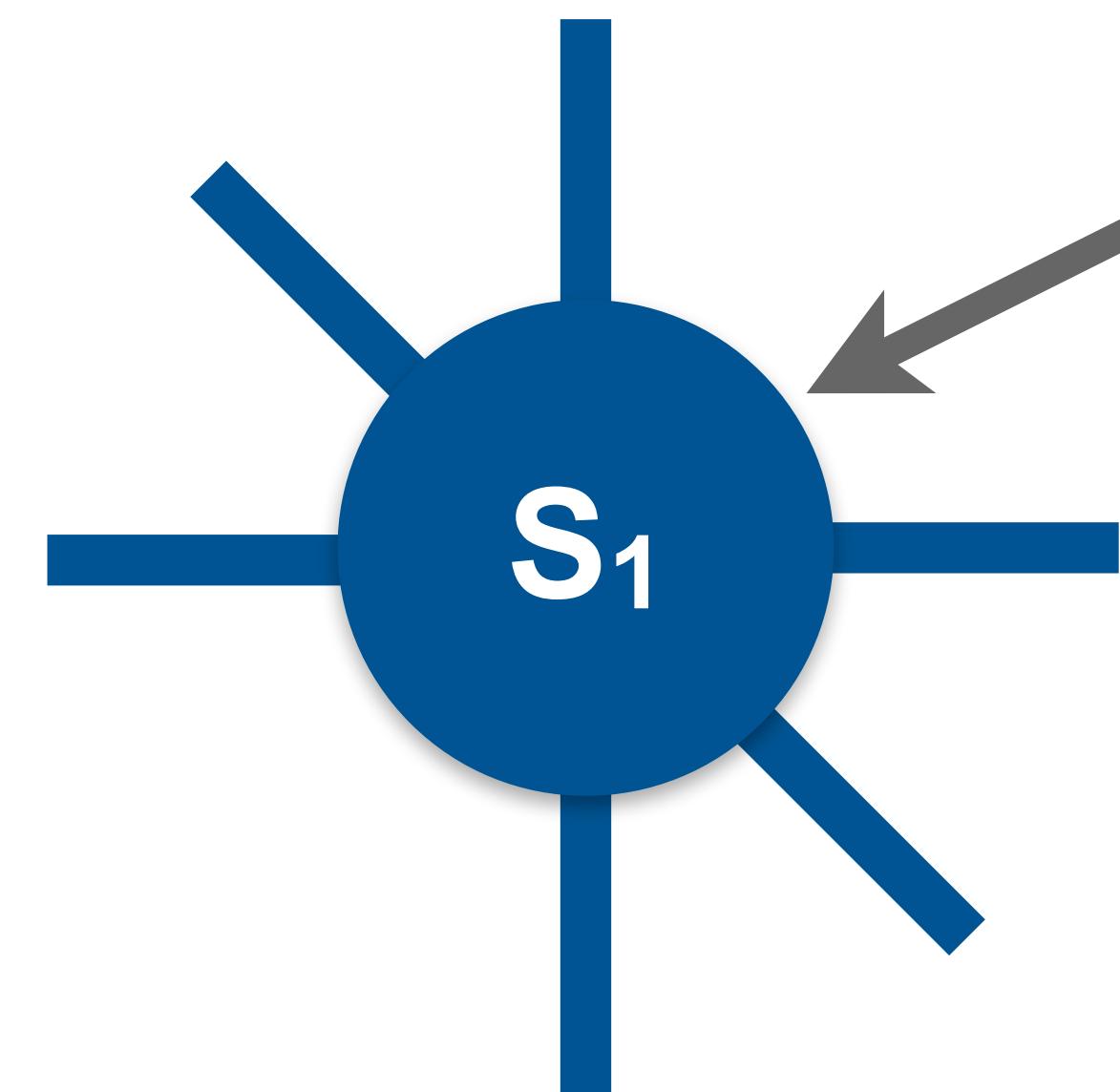
\mathcal{C}_5

Challenges - Other Graphs

Source Graph



Target Graph

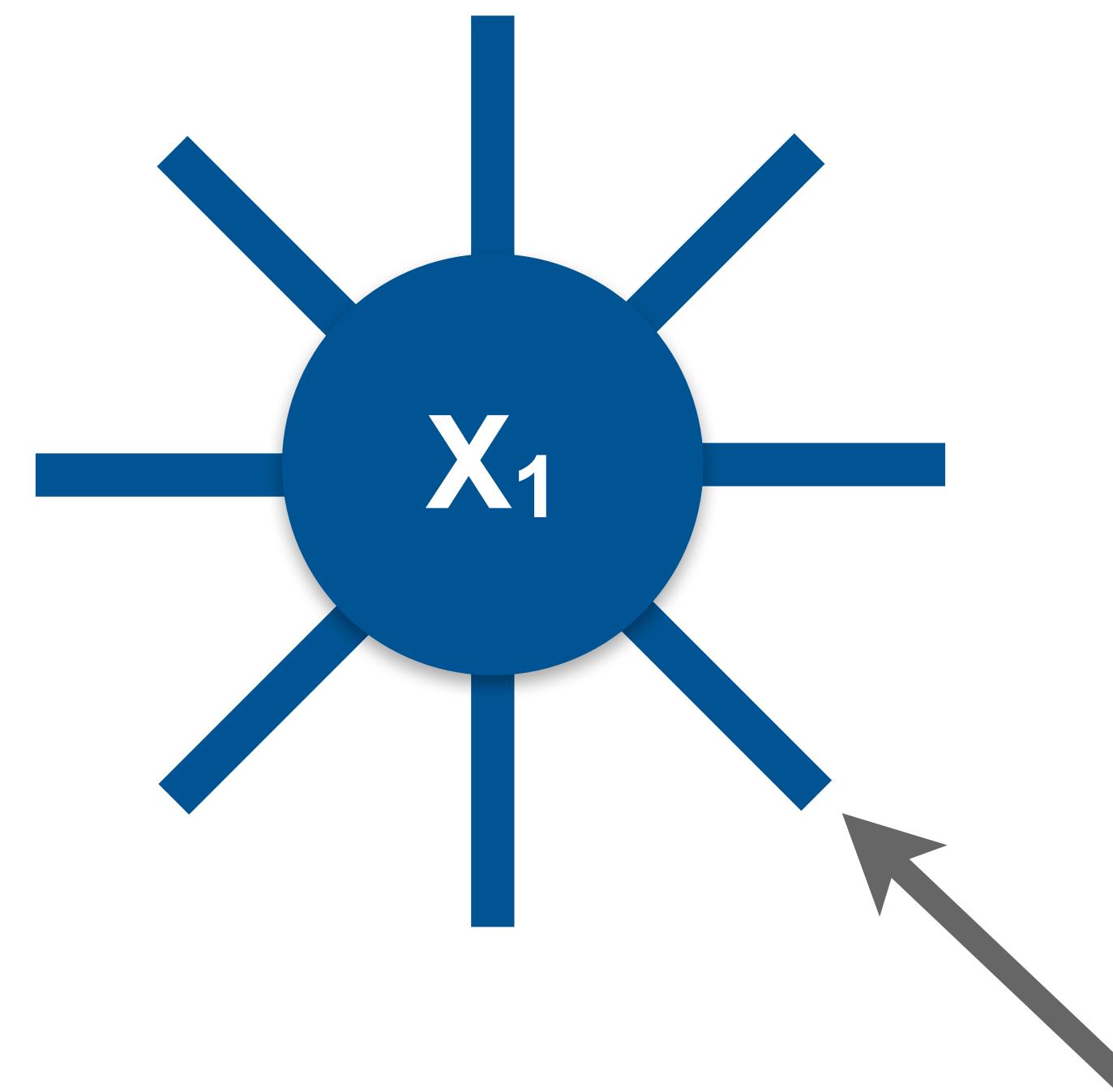


only 6
edges!

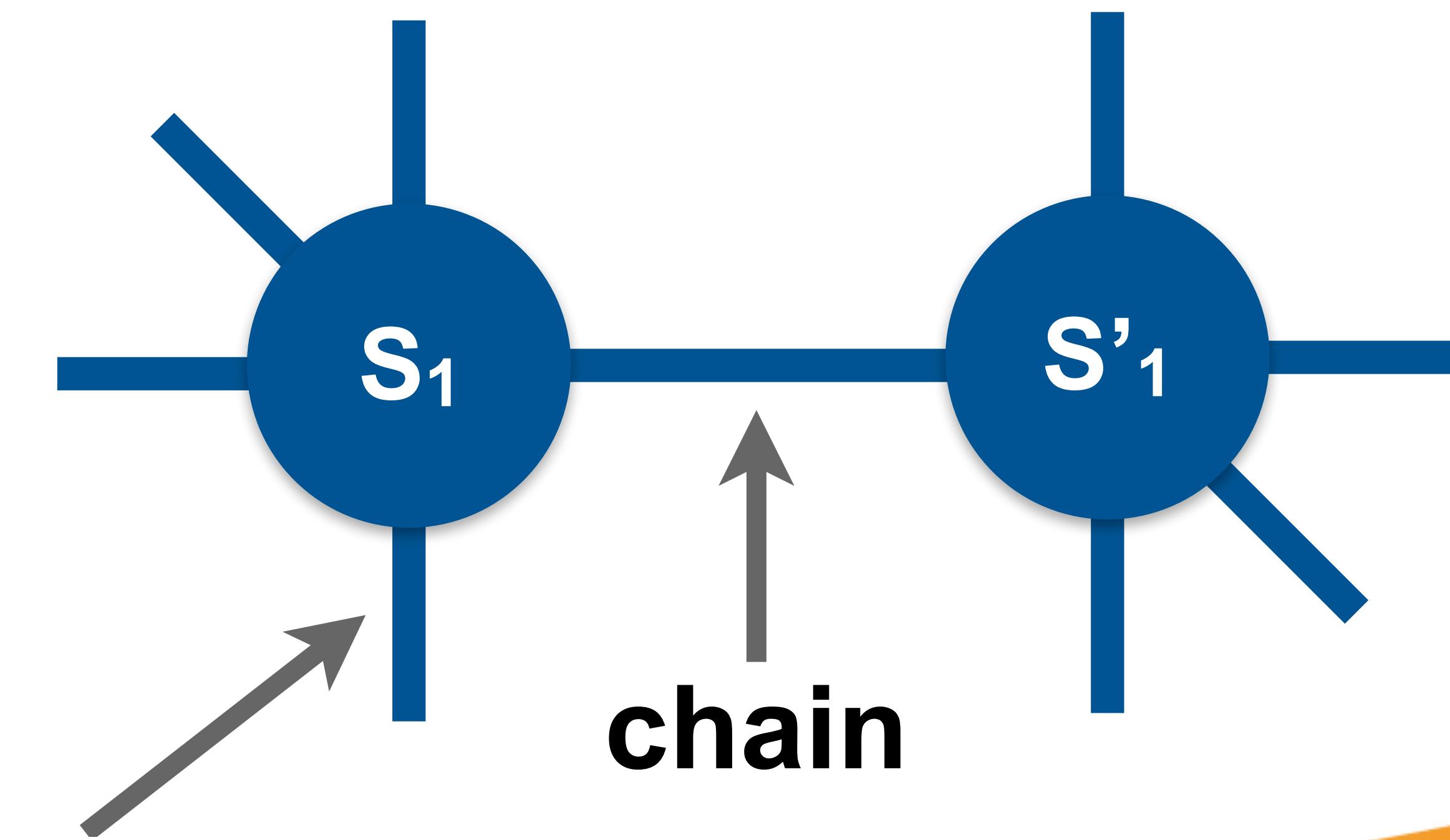
Challenges - Other Graphs

Warning: Can for worms!

Source Graph



Target Graph



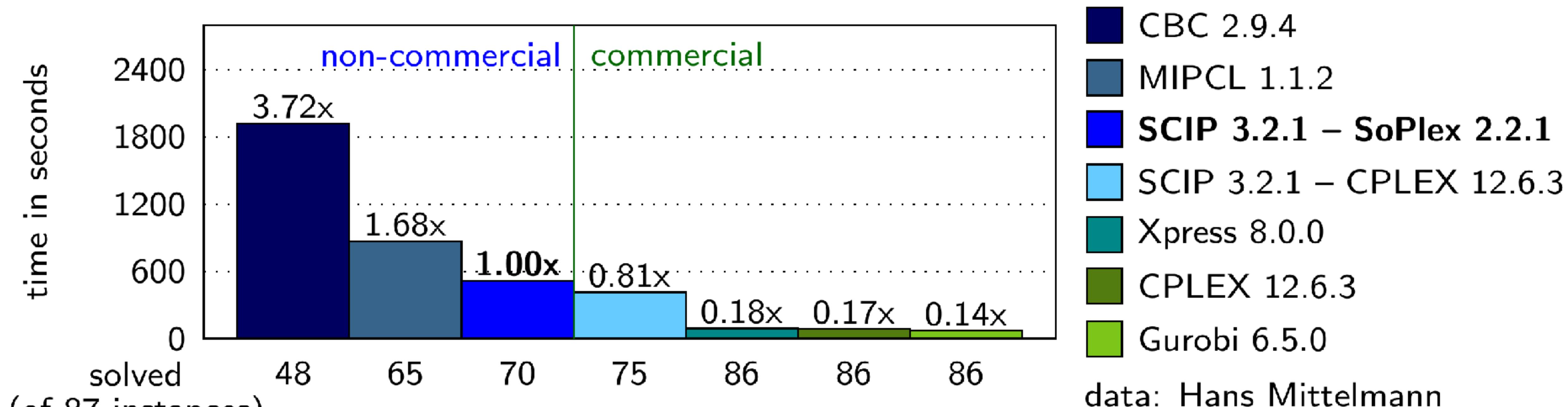
logical interactions

Challenges (second-order approx.)

- Low-temp sampler, not adiabatic optimizer
 - Analog noise
 - Biases
 - Programming errors
 - et. al. (see QPU guide)

“Technical Description of the D-Wave Quantum Processing Unit”

Despite Challenges...



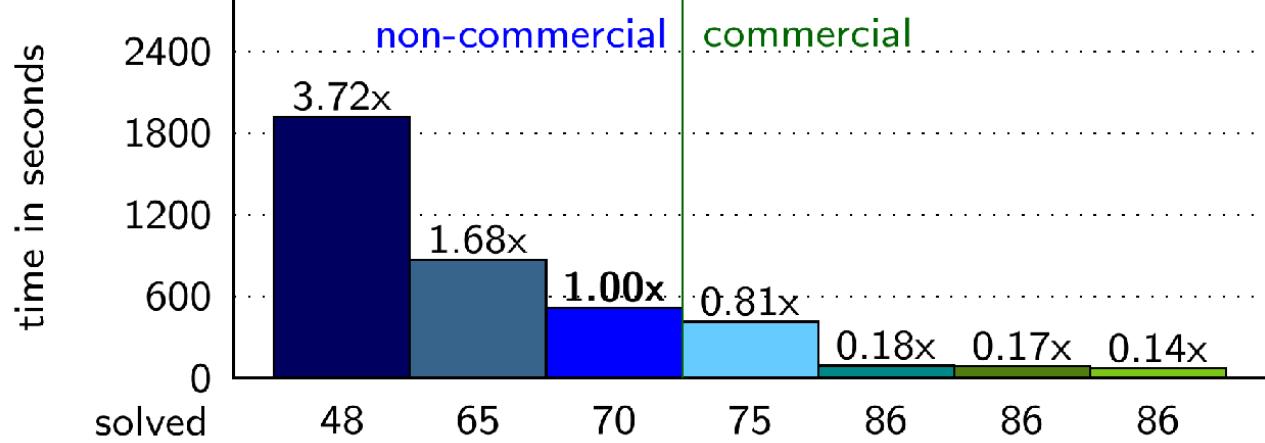
D:wave

Benchmarking a D-Wave

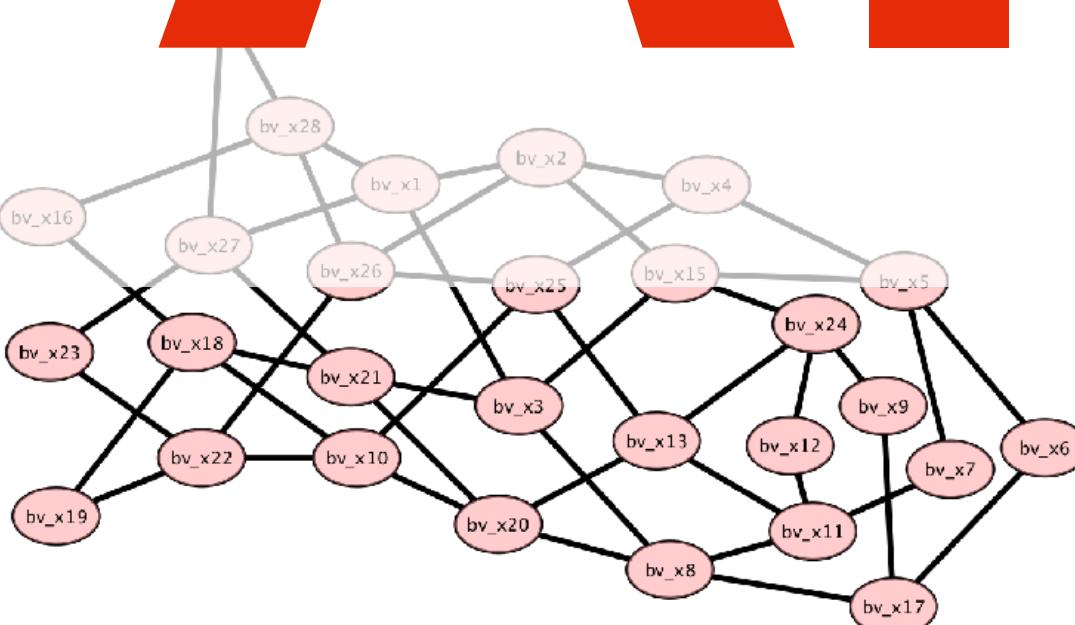
The Benchmarking Problem

ISTI '16


QPLIB
DIMACS
Max-Clique

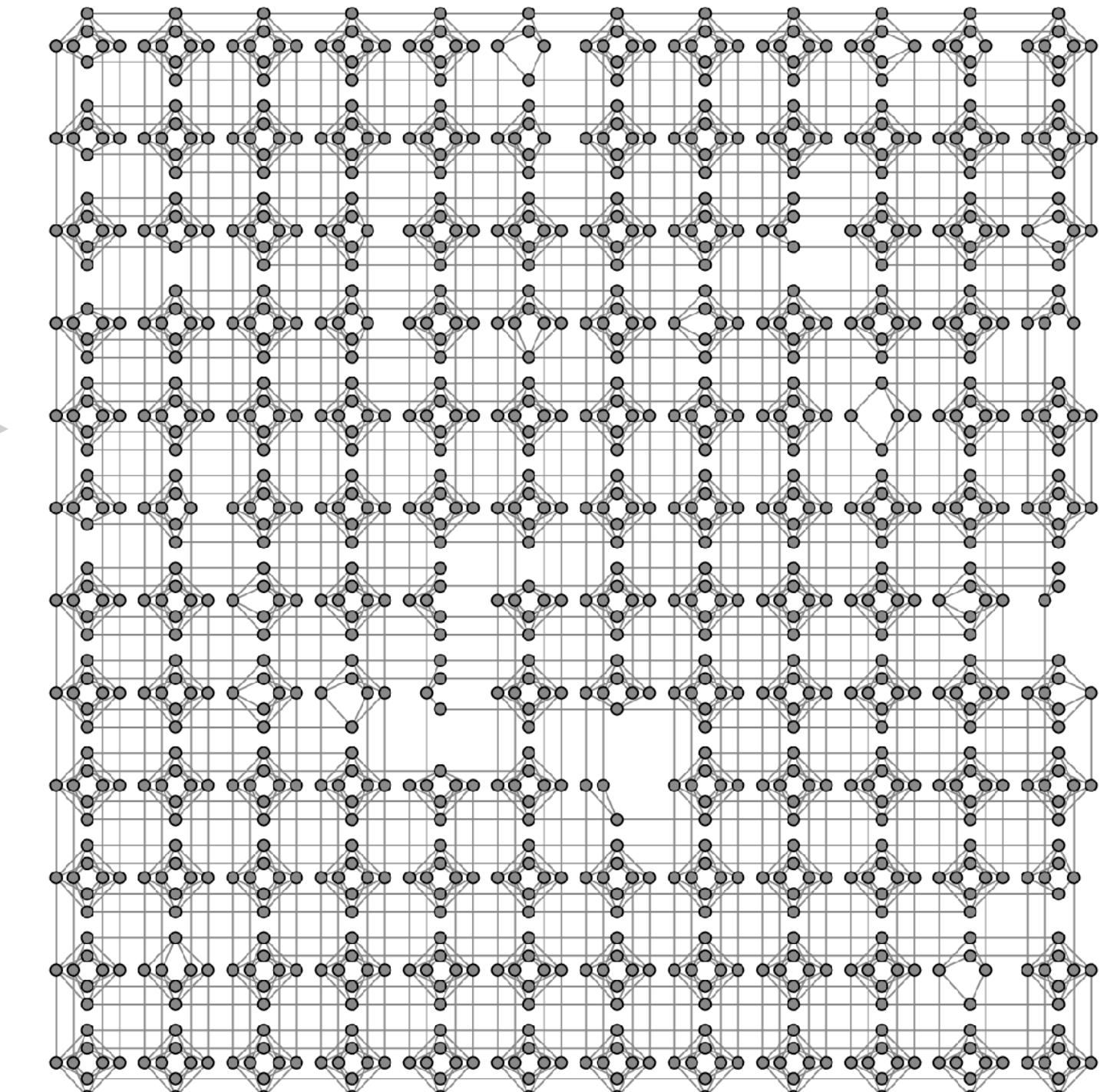






<https://youtu.be/uEsfVAVd5ks>





C_{12}

DW2X

A Benchmarking Stopgap

B-QP
Solvers

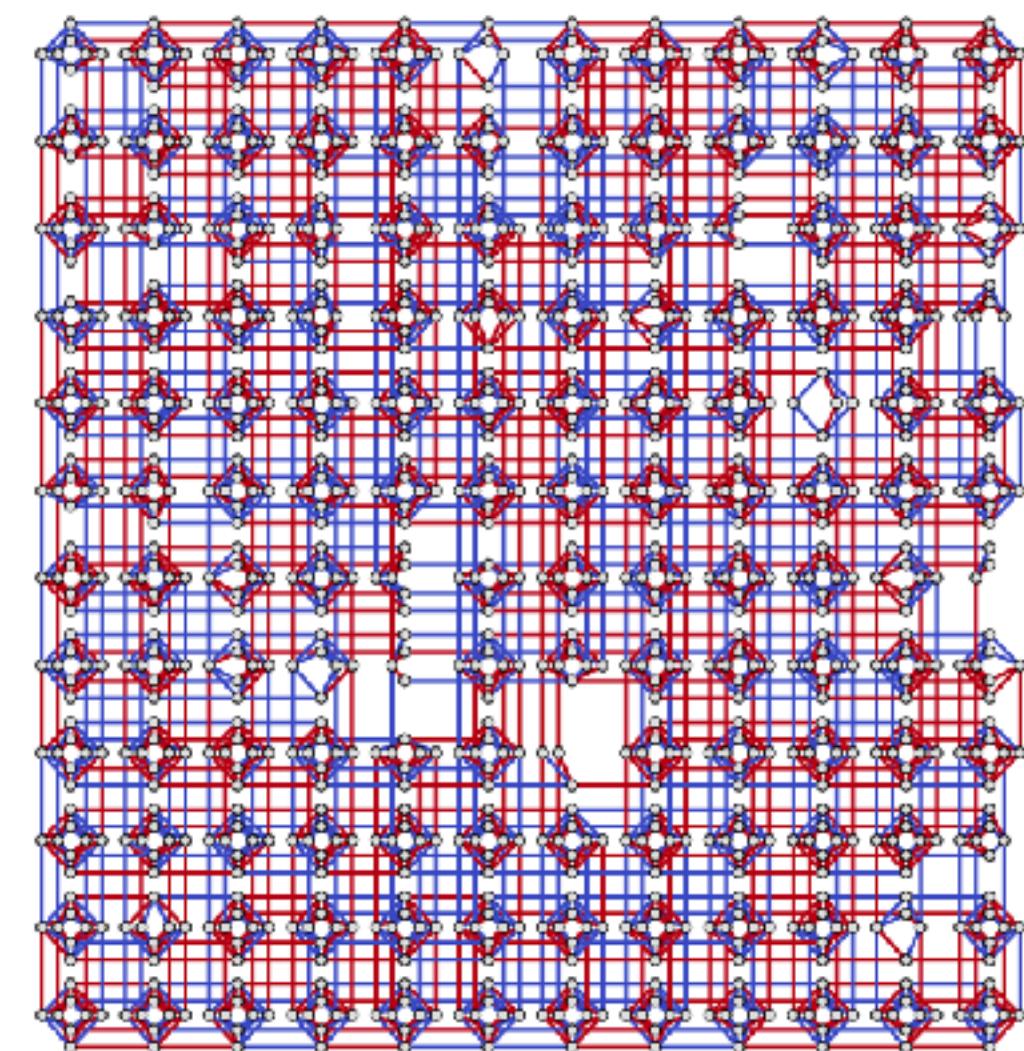


Google
or-tools

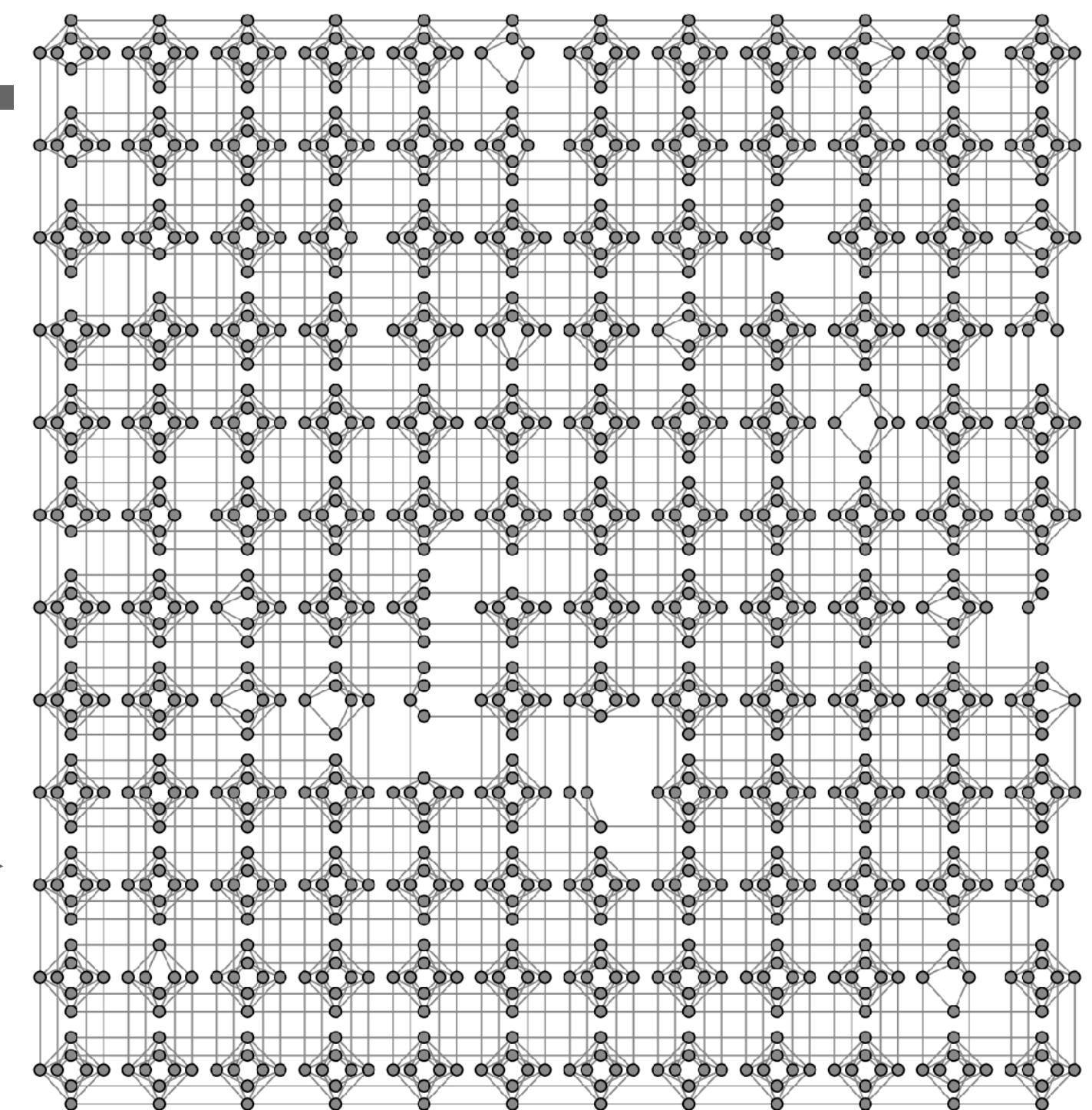


...
Zn

Problem
Generation



D-Wave



C₁₂ DW2X

Problems with Problem Generation

- How hard are randomly generated problems?
- Lessons learned from Random SAT

Hard and Easy Distributions of SAT Problems

David Mitchell

Dept. of Computing Science
Simon Fraser University

Bart Selman

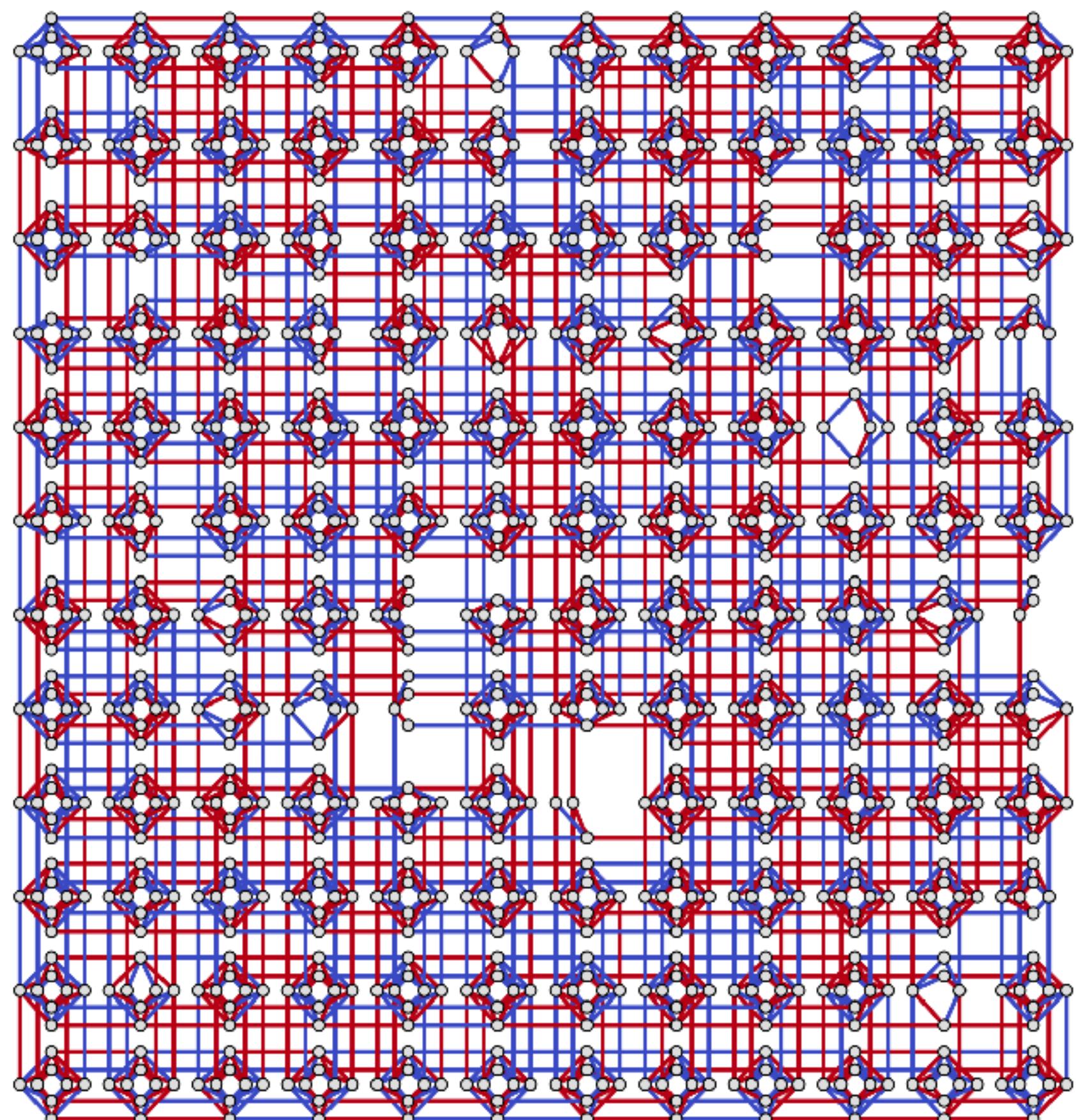
AT&T Bell Laboratories
Murray Hill, NJ 07974

Hector Levesque*

Dept. of Computer Science
University of Toronto

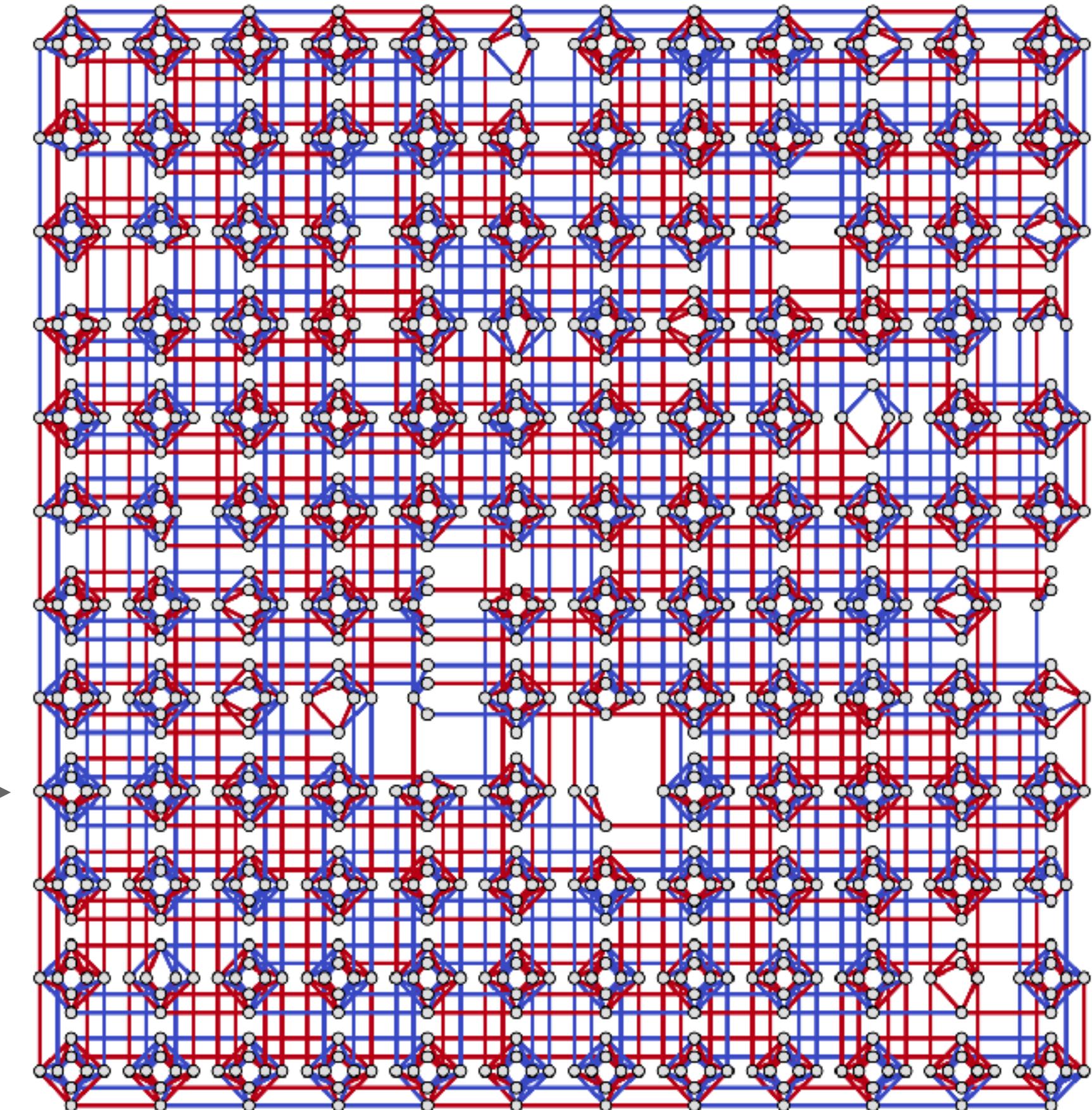
AAAI-92

Example: What's the Difference?



HARDEST
super frustrated
system

EASIEST
ferromagnet
in disguise!



The Key Challenge

How to generate a **HARD** D-Wave case

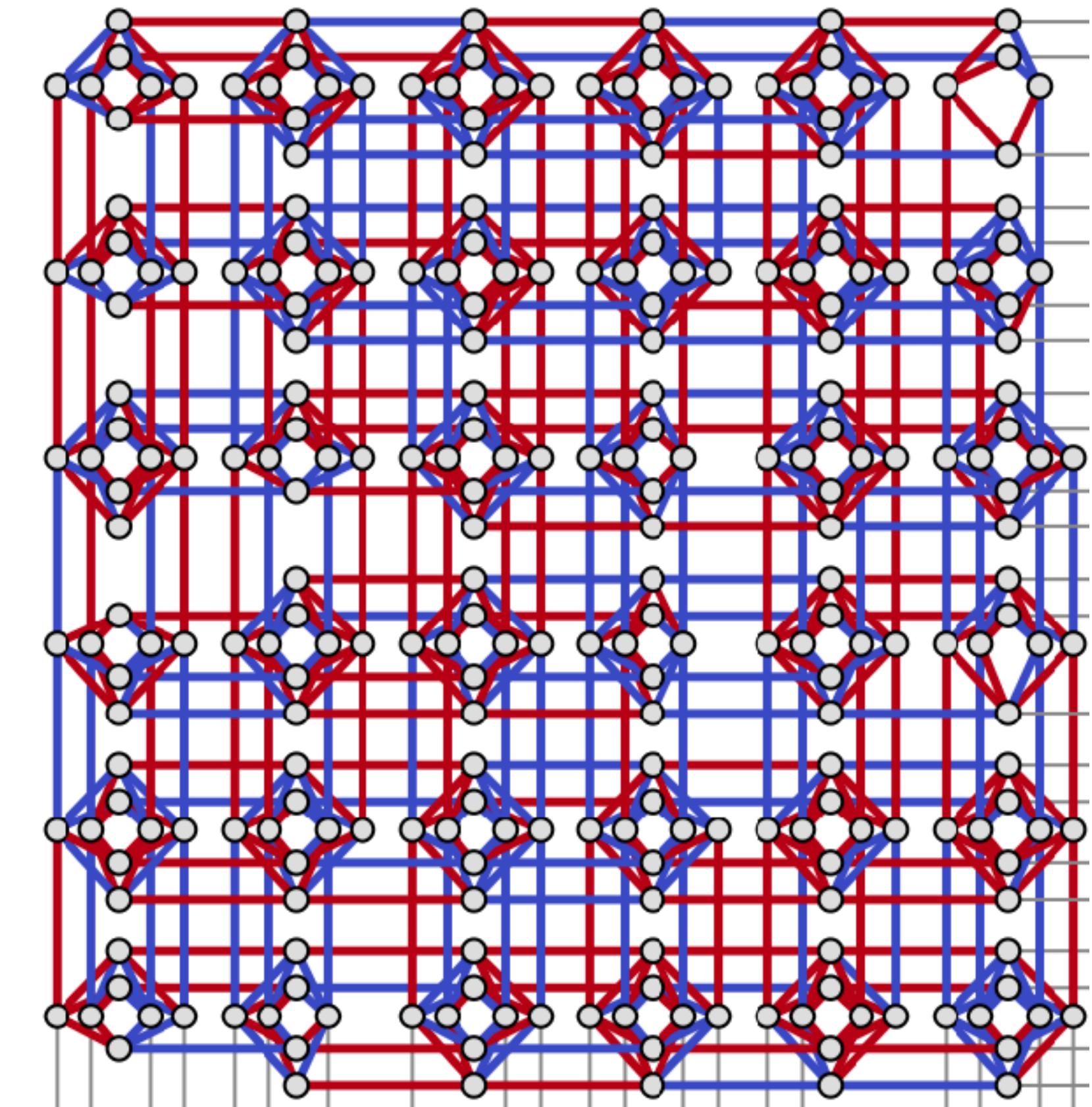


look to the literature

Proposed Problem Generators

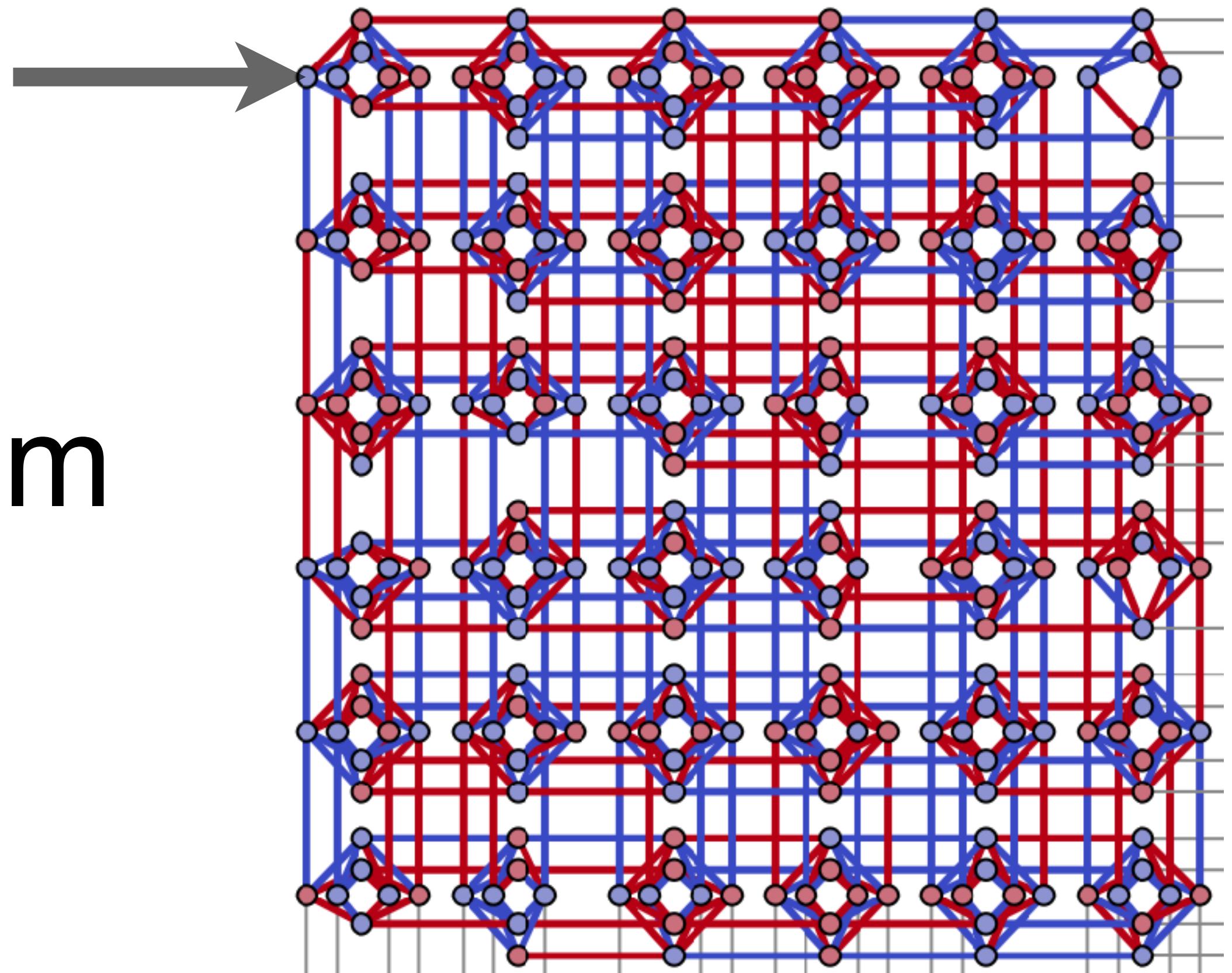
- **RAN-k**
 - set field to zero
 - set couplers at random
 - use k discrete steps

<https://arxiv.org/abs/1508.05087>



Proposed Problem Generators

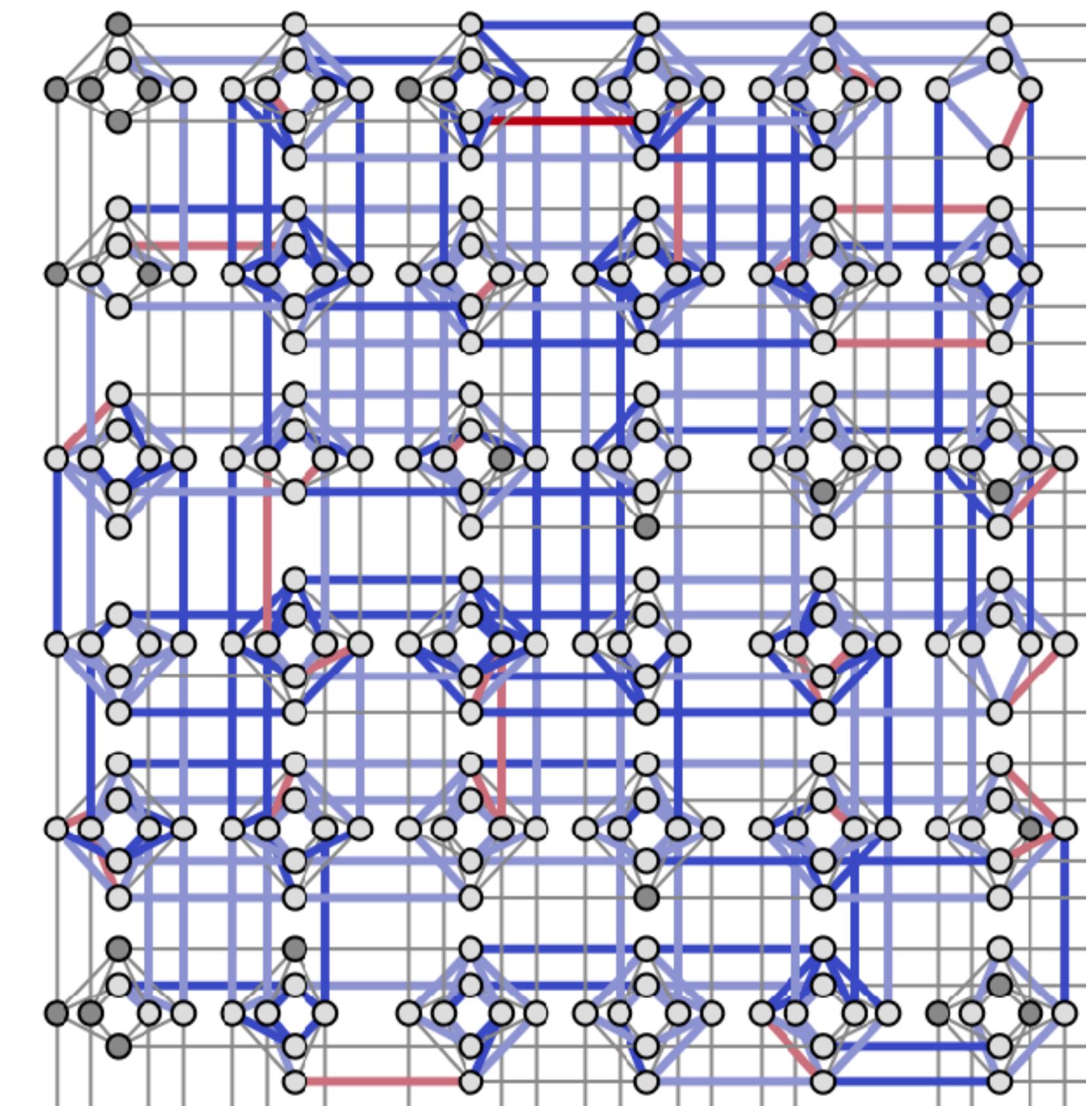
- **RANF-k**
 - set field at random
 - set couplers at random
 - use k discrete steps



Proposed Problem Generators

- Frustrated Loops (**FL**)
 - find random cycles
 - add one edge of frustration
 - overlay multiple cycles

<https://arxiv.org/abs/1508.05087>

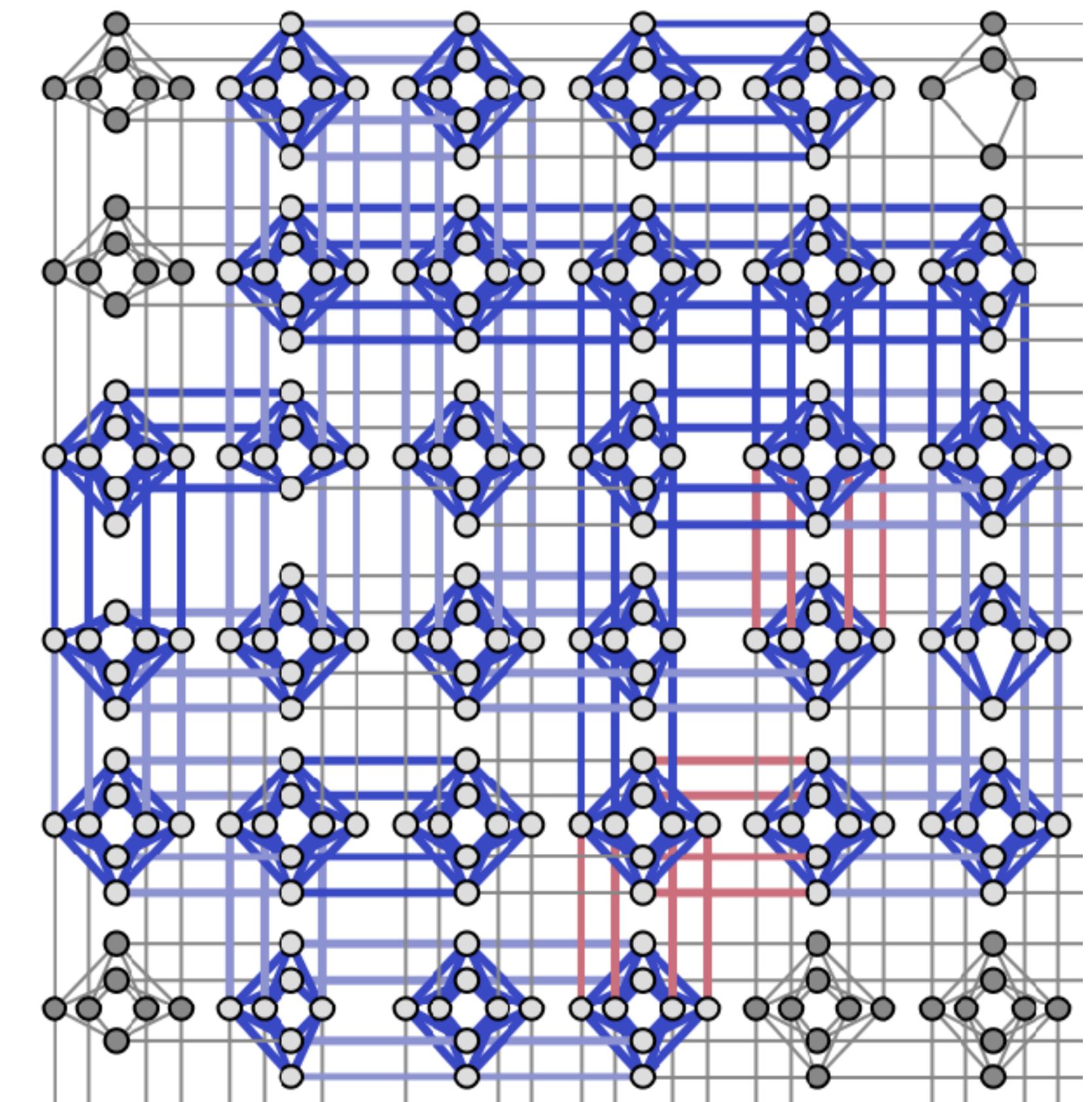


Proposed Problem Generators

- Frustrated Cluster Loops (FCL)
 - find random cycles
 - add one edge of frustration
 - overlay multiple cycles

<https://arxiv.org/abs/1701.04579>

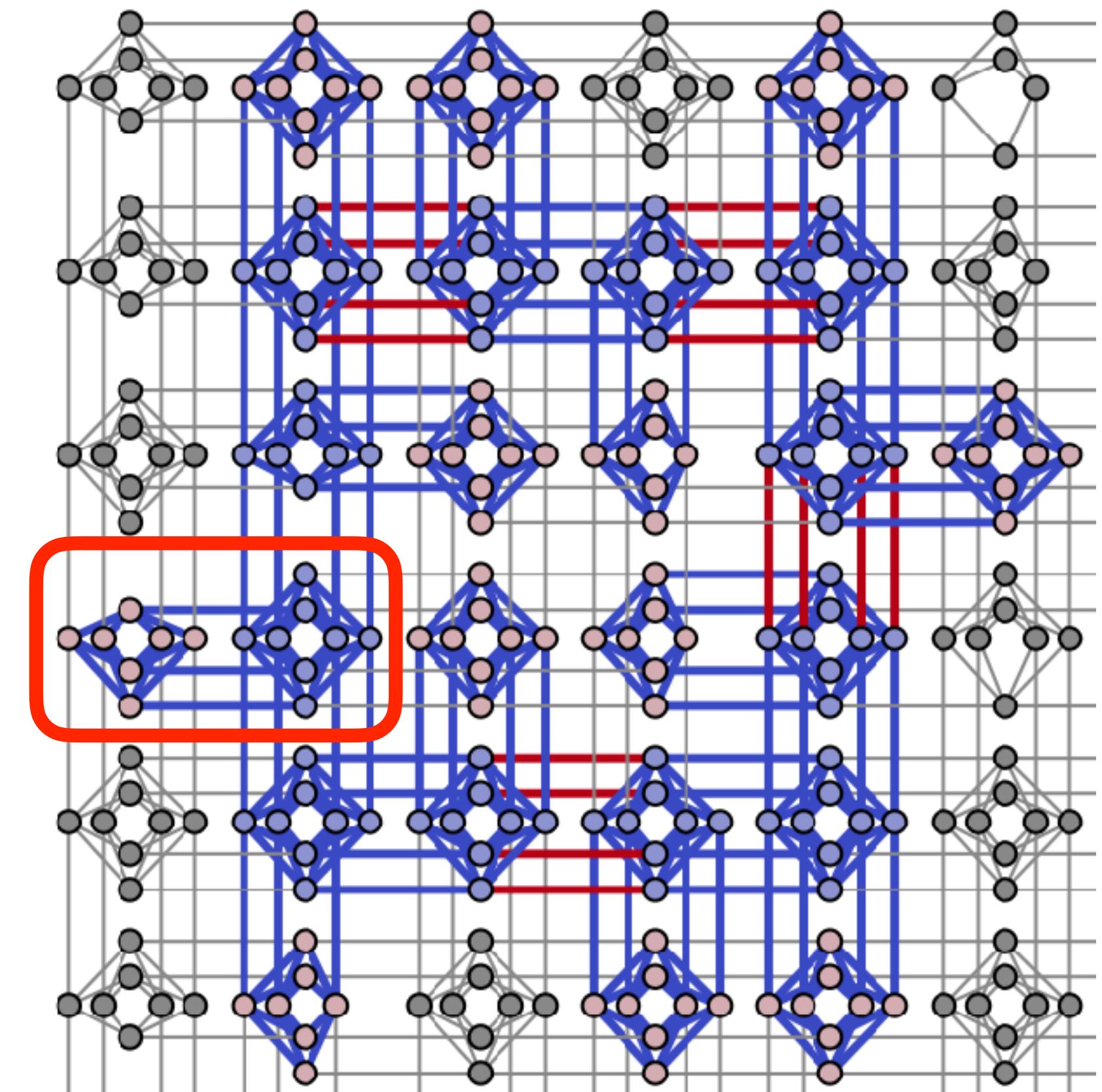
<https://arxiv.org/abs/1703.00622>



Proposed Problem Generators

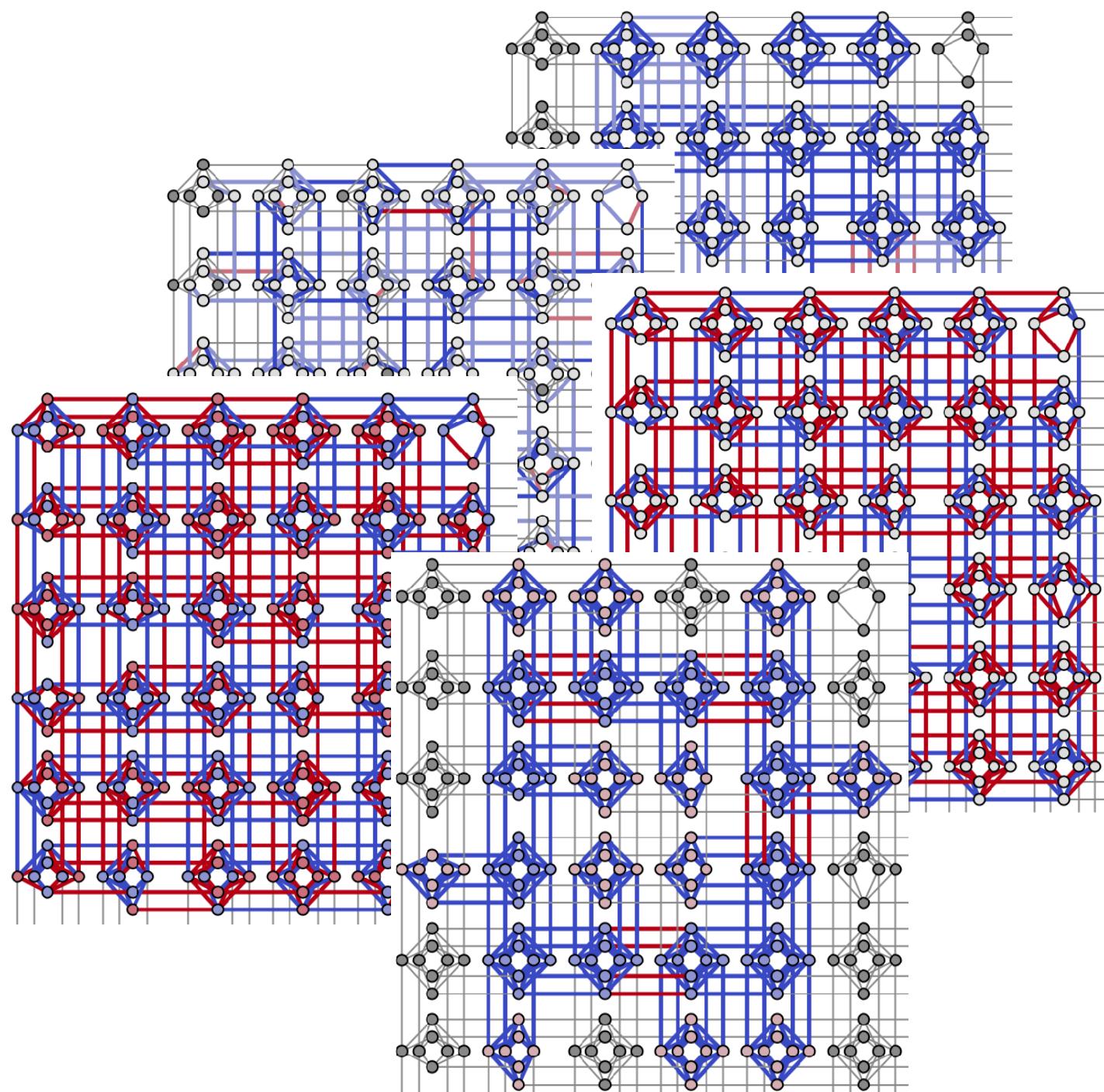
- Weak-Strong Cluster Networks (**WSCN**)
- random grid of two cell gadgets

<https://arxiv.org/abs/1512.02206>



What to Compare?

Cases



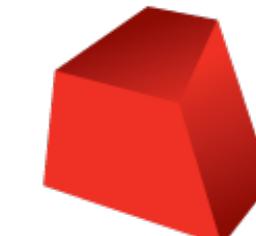
?

Solvers

qbsolv



D-wave



**GUROBI
OPTIMIZATION**

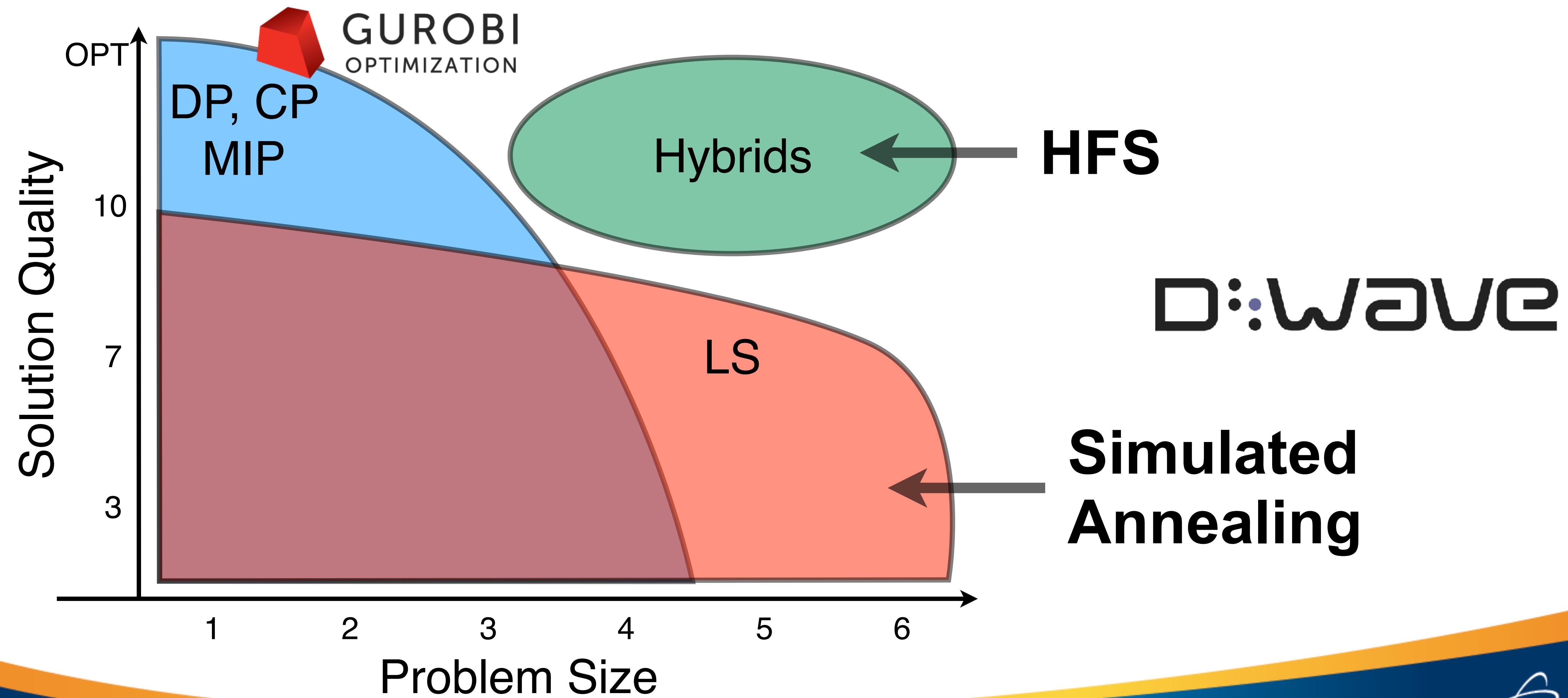
HFS

SA

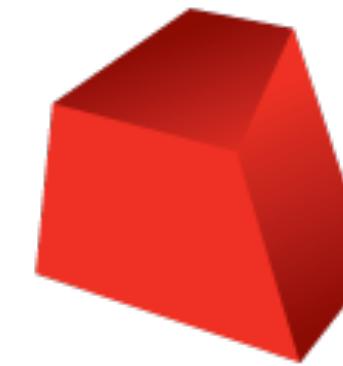


**Google
or-tools**

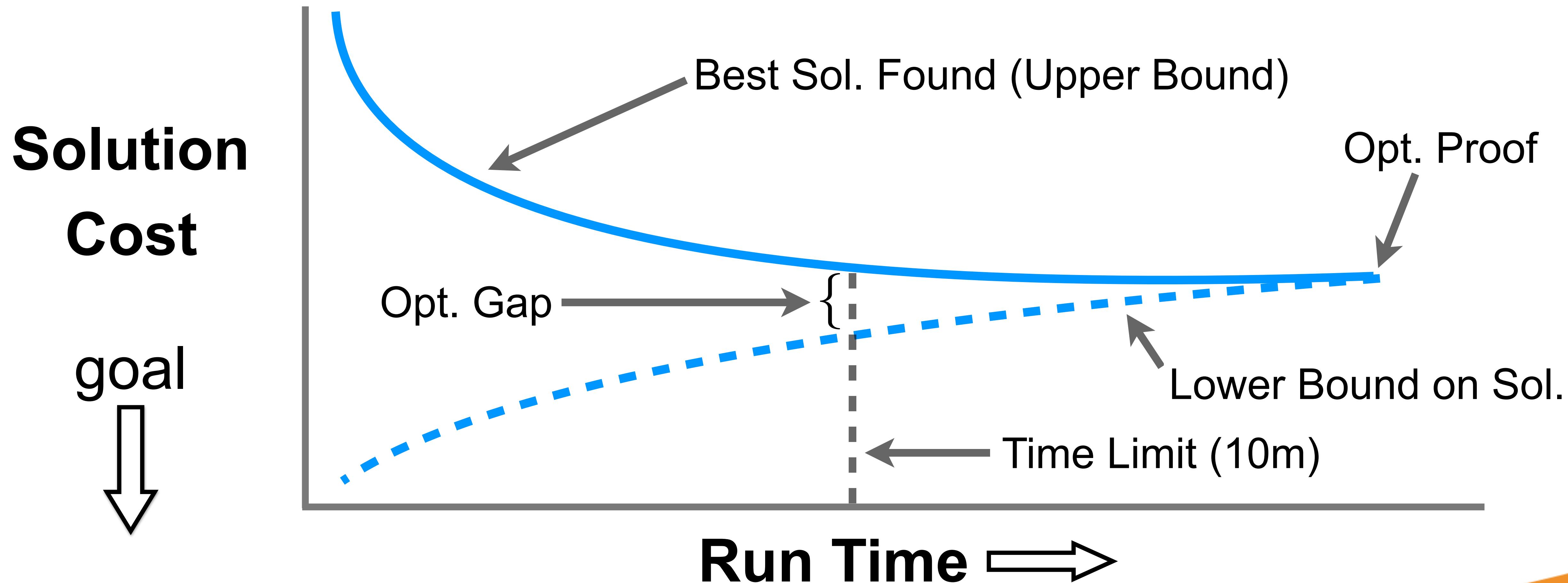
Solution Method Classes



Classic B-QP Solver



GUROBI
OPTIMIZATION



HFS Solver

Hamze-de Freitas-Selby (HFS)

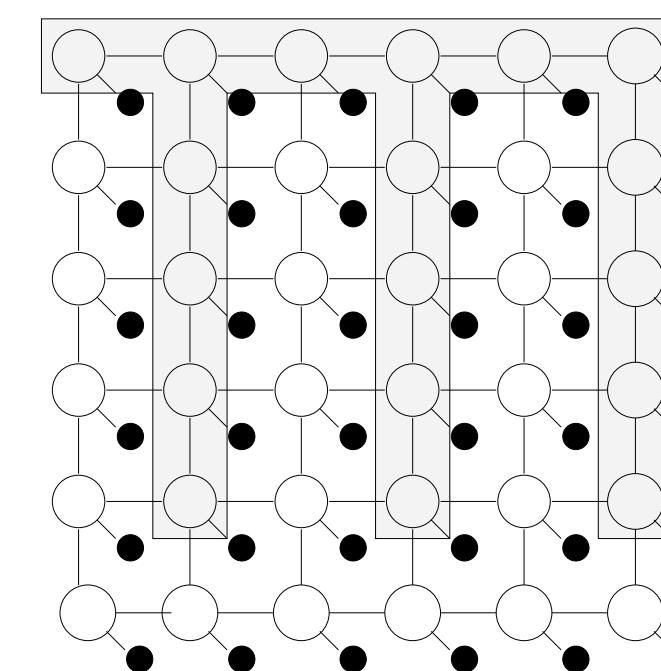
<https://arxiv.org/pdf/1207.4149.pdf>

Low Treewidth

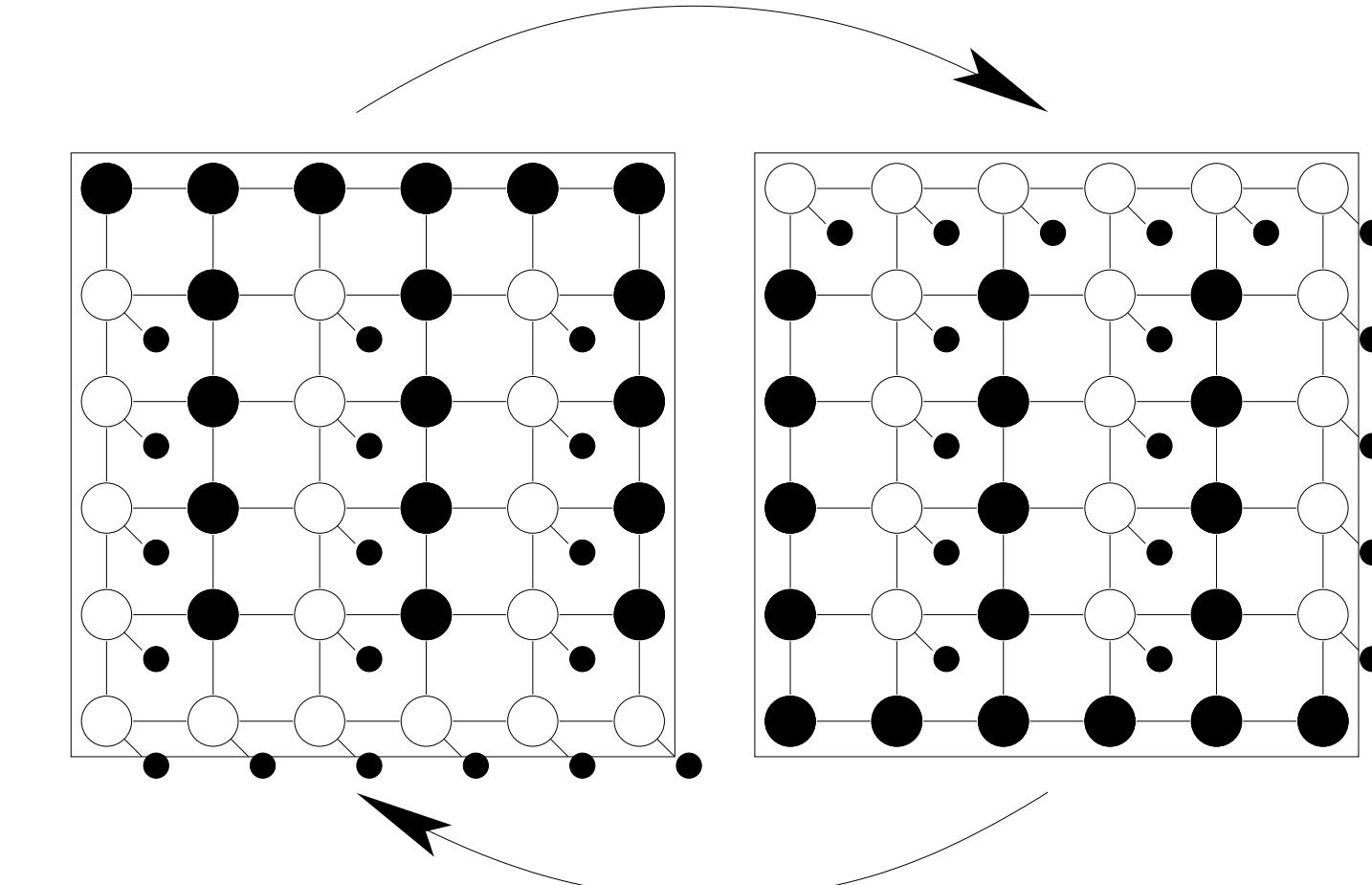
<https://github.com/alex1770/QUBO-Chimera>

HF
S

Subgraphs



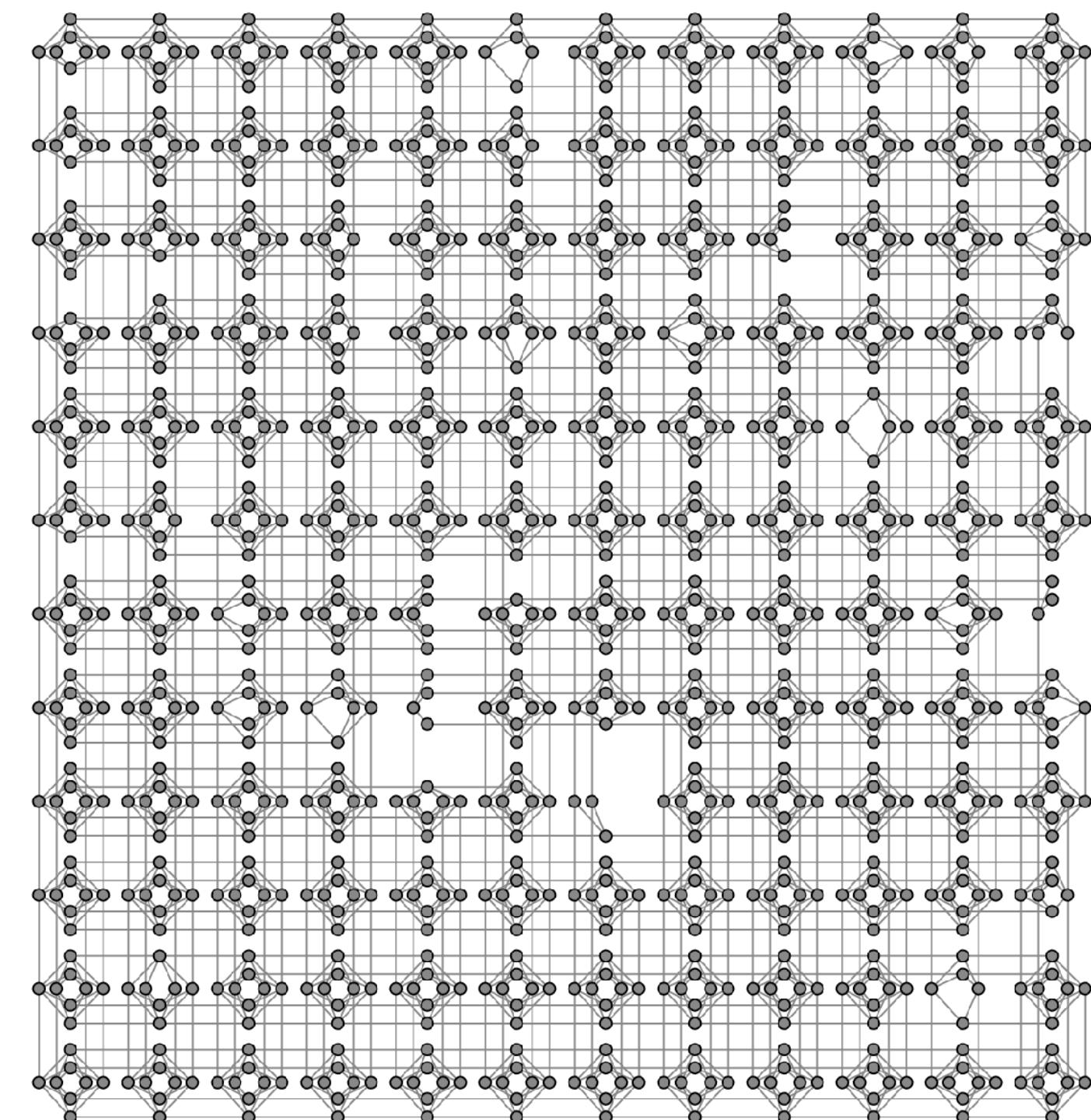
Optimization Loop using DP



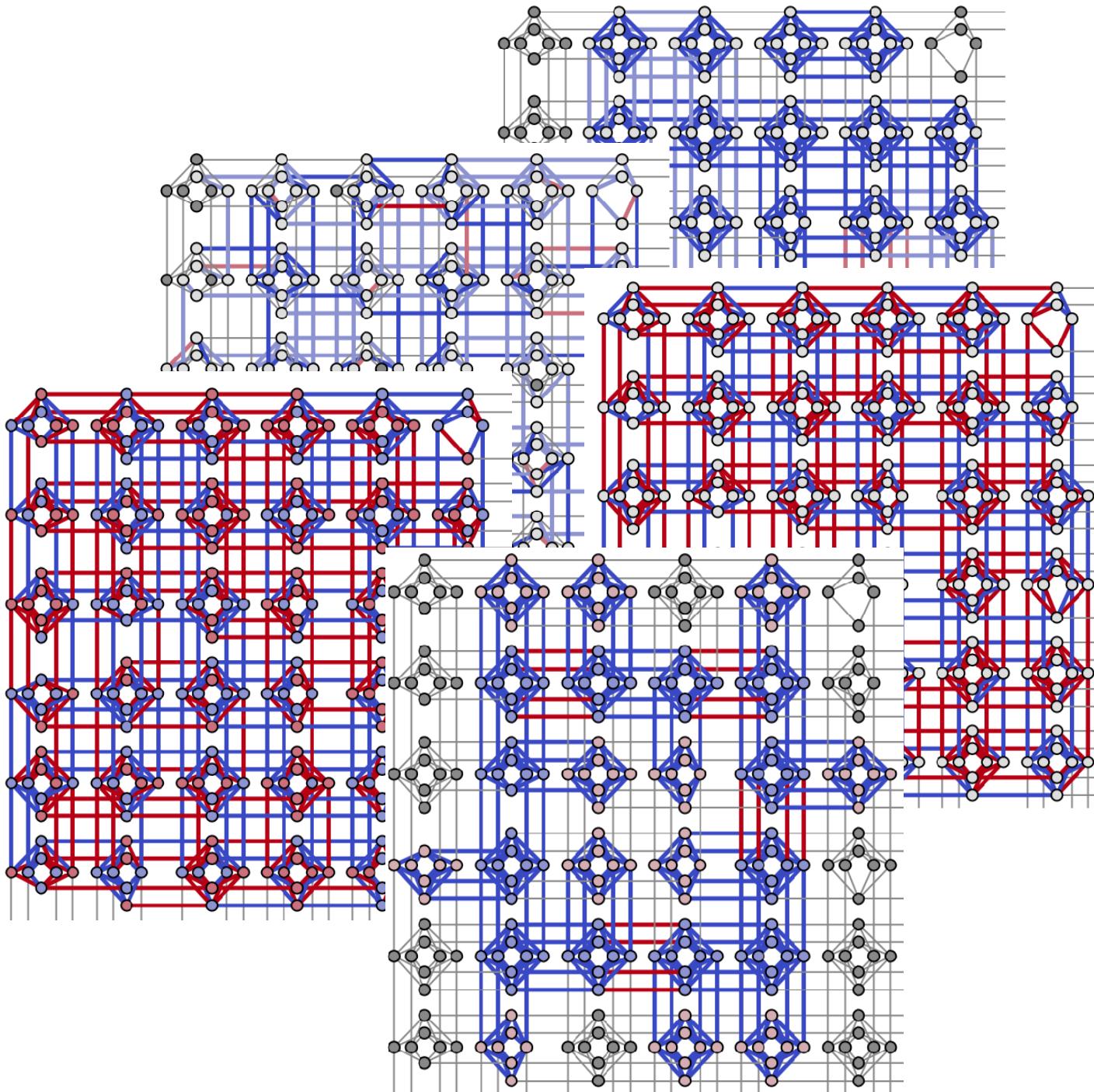
D-Wave For Optimization

- 10,000 samples @ 5 micro seconds each
- Random gauge transform every 100 samples
- Use the **best** of all samples
- Takes 3.5 seconds on the QPU

D:wave



D-Wave Instance Generator (DWIG)



BQPSOLVERS

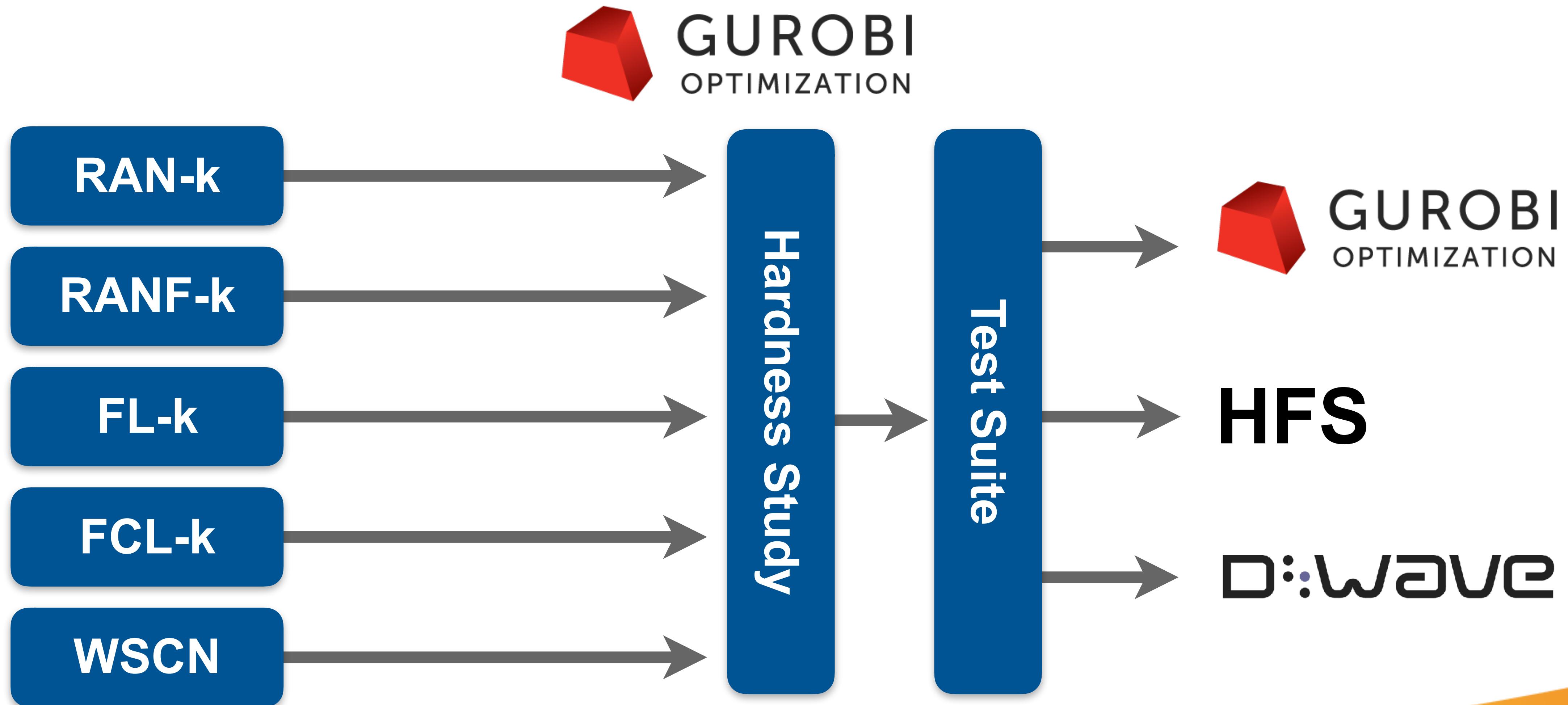
D:wave

HFS



Time for the Fun Stuff!

Structure of the Benchmark Study



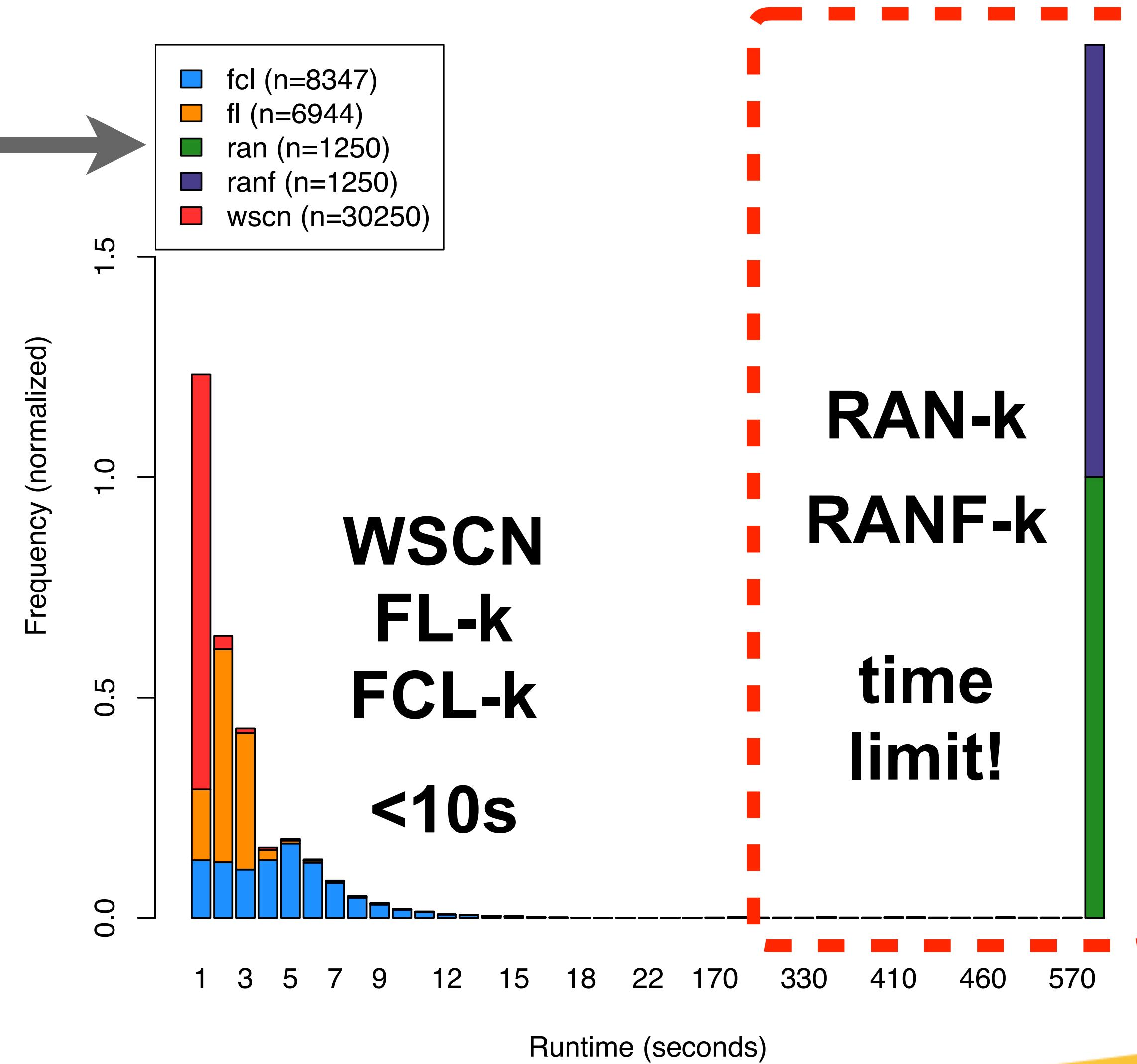
Problem Hardness Test

**Runtime Distribution
Color Coded
by Problem**

40,000+
Cases

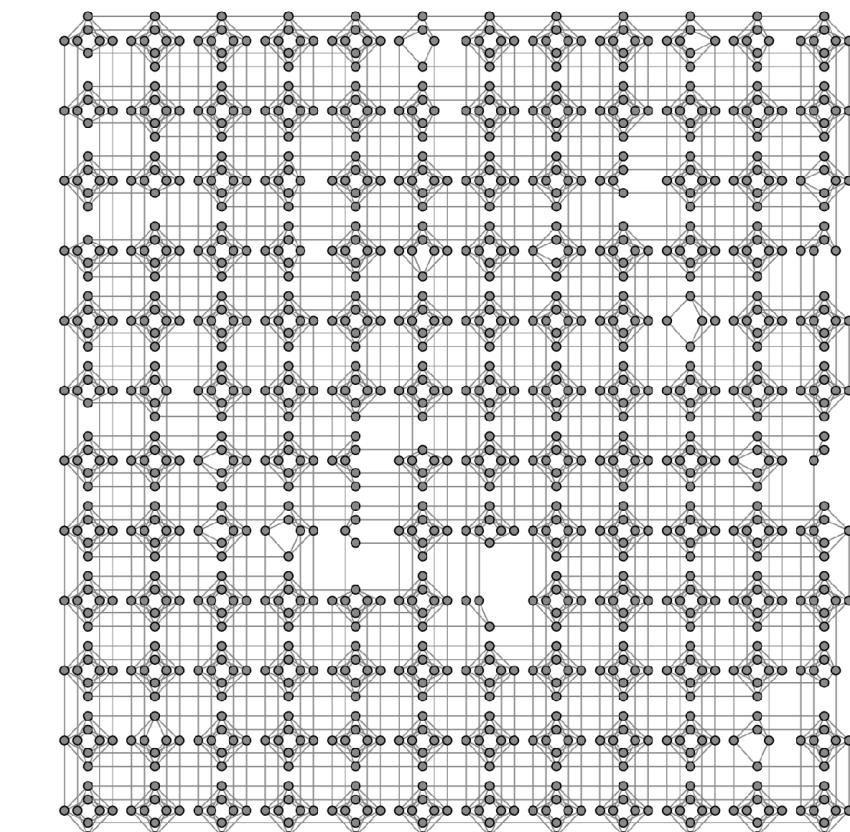
What about the
values of k?

$$2^{1000} \approx 10^{300}$$



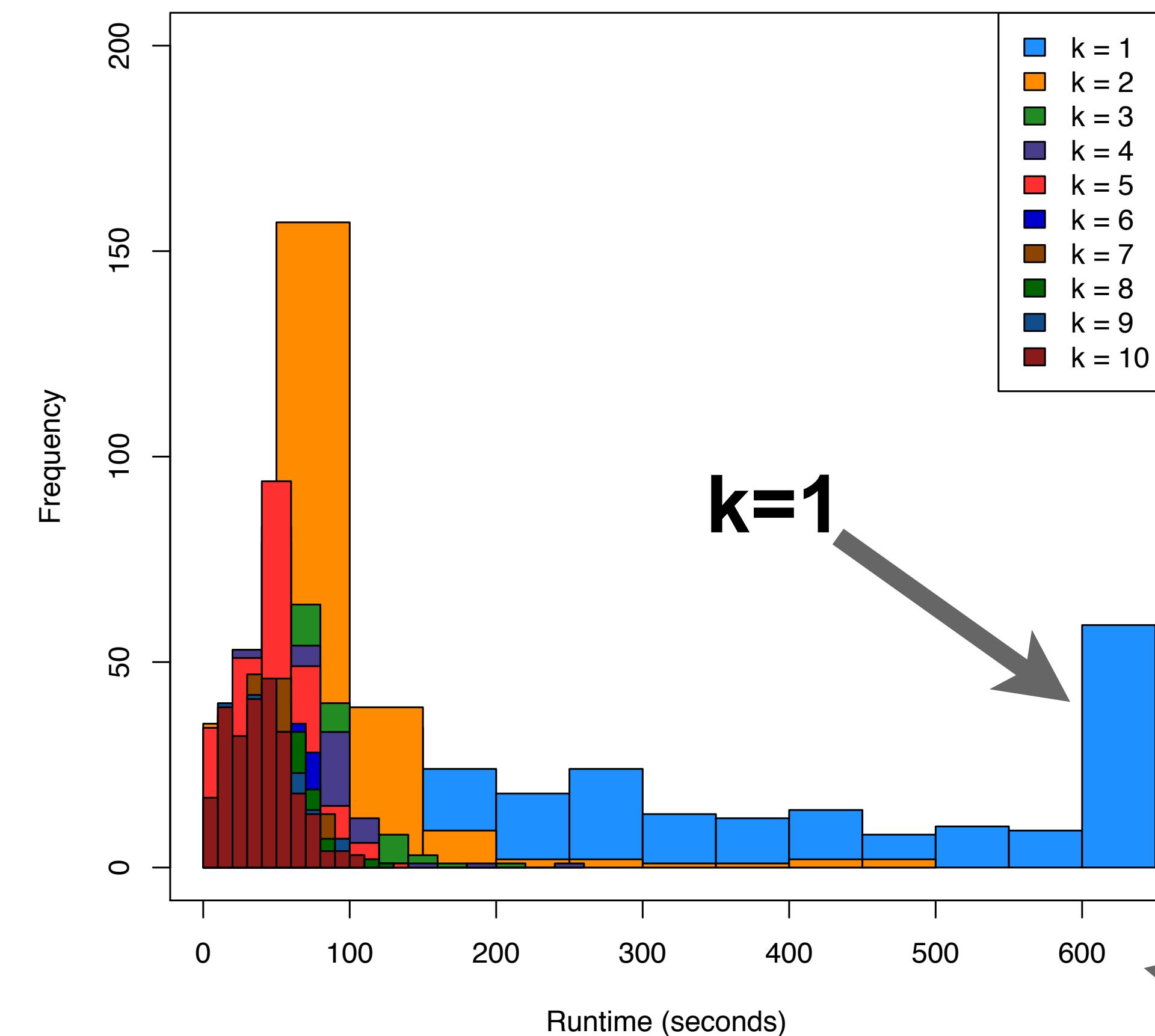
600 second
time limit

C_{12} Max Size

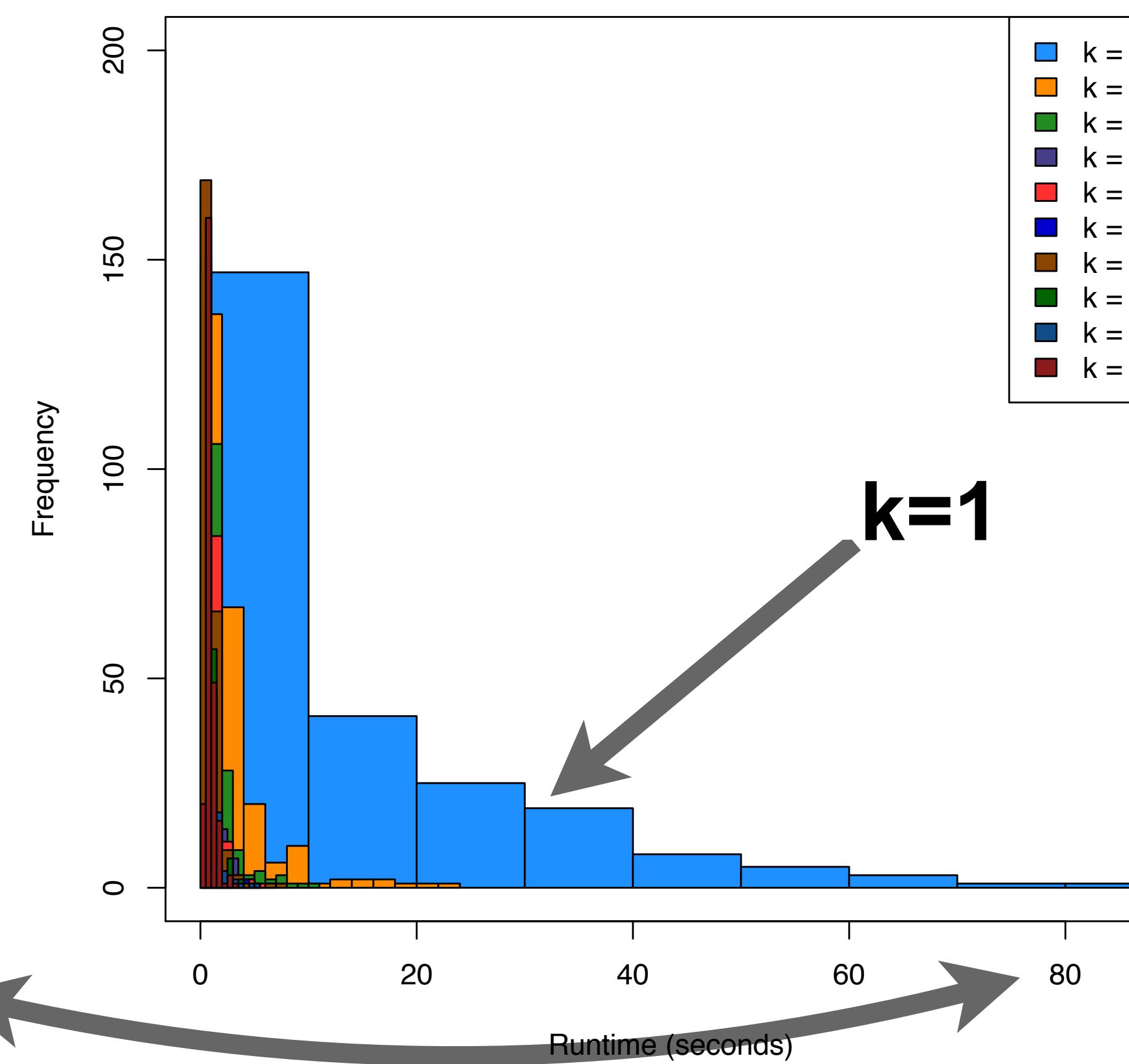


How does Hardness change with K?

RAN-k

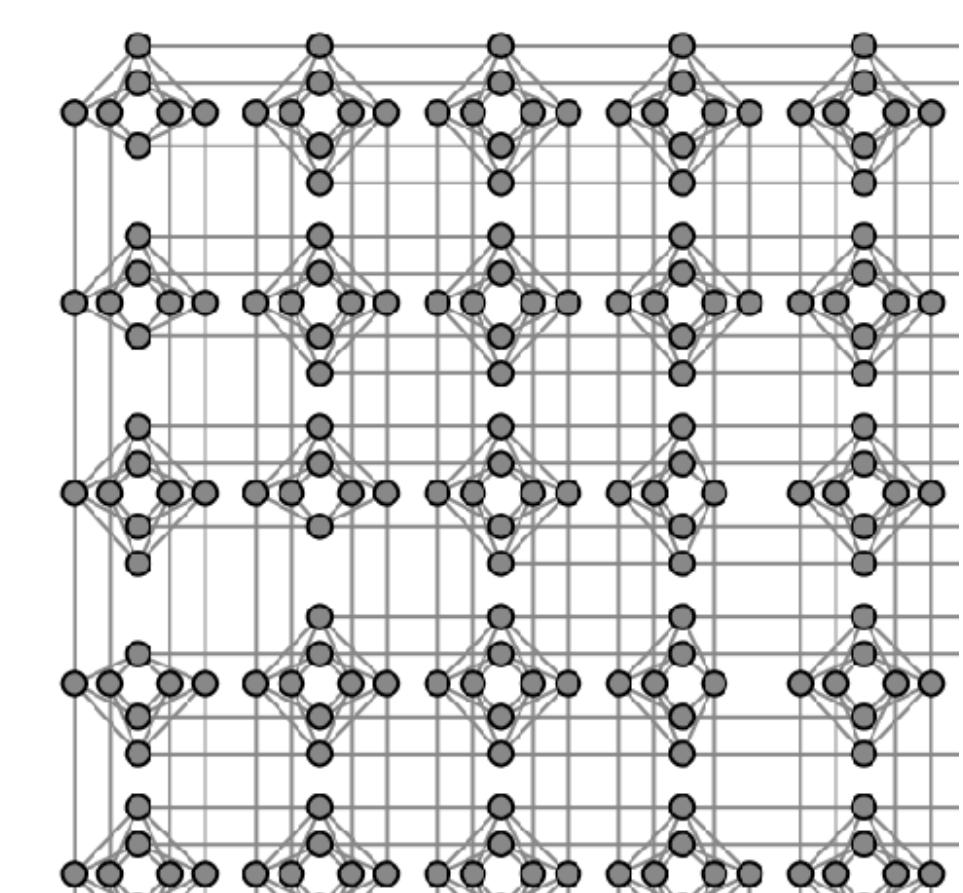


RANF-k



RAN > RANF

GUROBI
OPTIMIZATION
optimality proof in
600 sec. or less



$$C_5 \approx 2^{194} \approx 10^{58}$$

These guys know what they are talking about

“Have you tried setting the couplers to -1,1 at random?”

09/2016



Marc
(T-4)



Andrey
(CNLS/T-4)



Sidhant
(T-5)

Focusing A Detailed Study of RAN-1



GUROBI
OPTIMIZATION

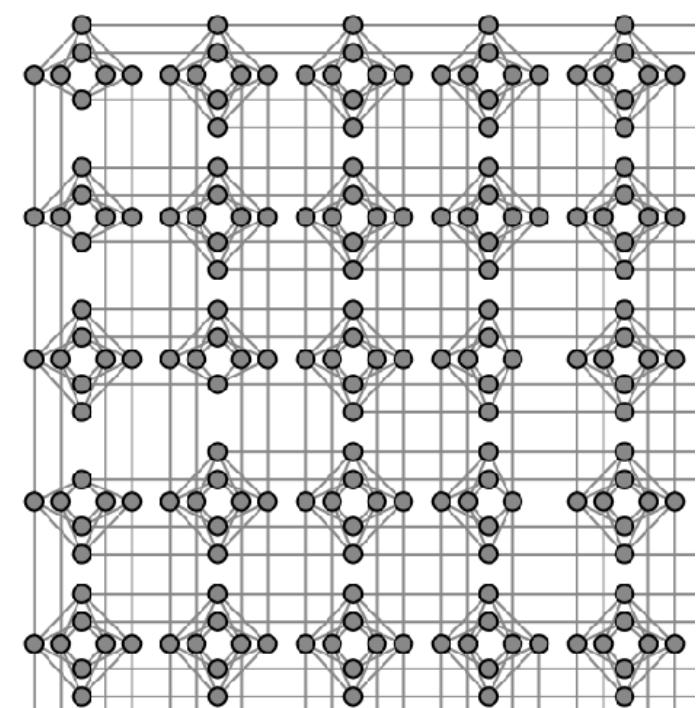
HFS

D-wave

(Similar story for RANF-1, omitted for time)

Detailed Runtime Studies

Part 1



C_5

Quality Validation
(Known Global Optimum)

Wishful Extrapolation



GUROBI
OPTIMIZATION



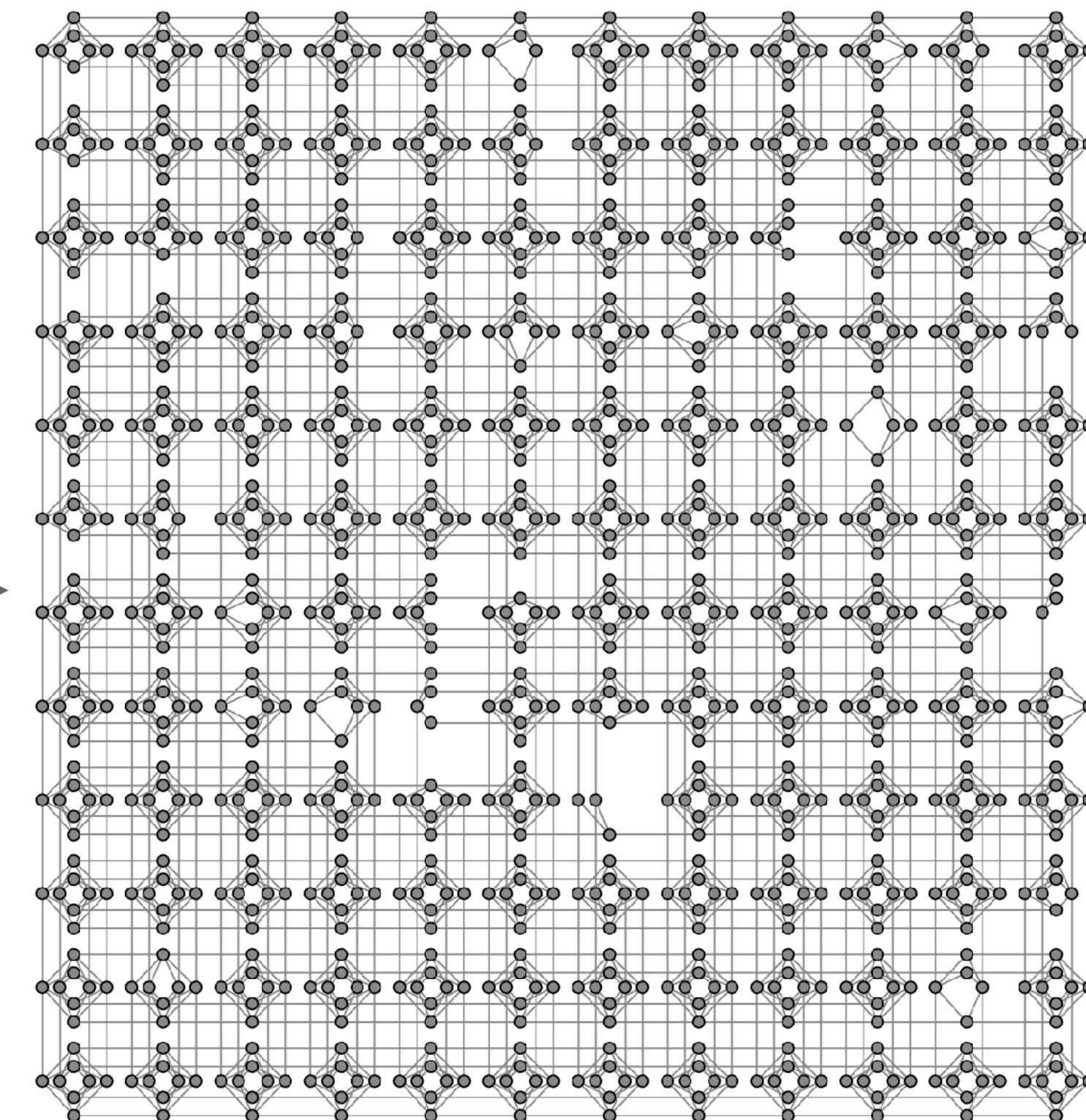
D-wave



HFS

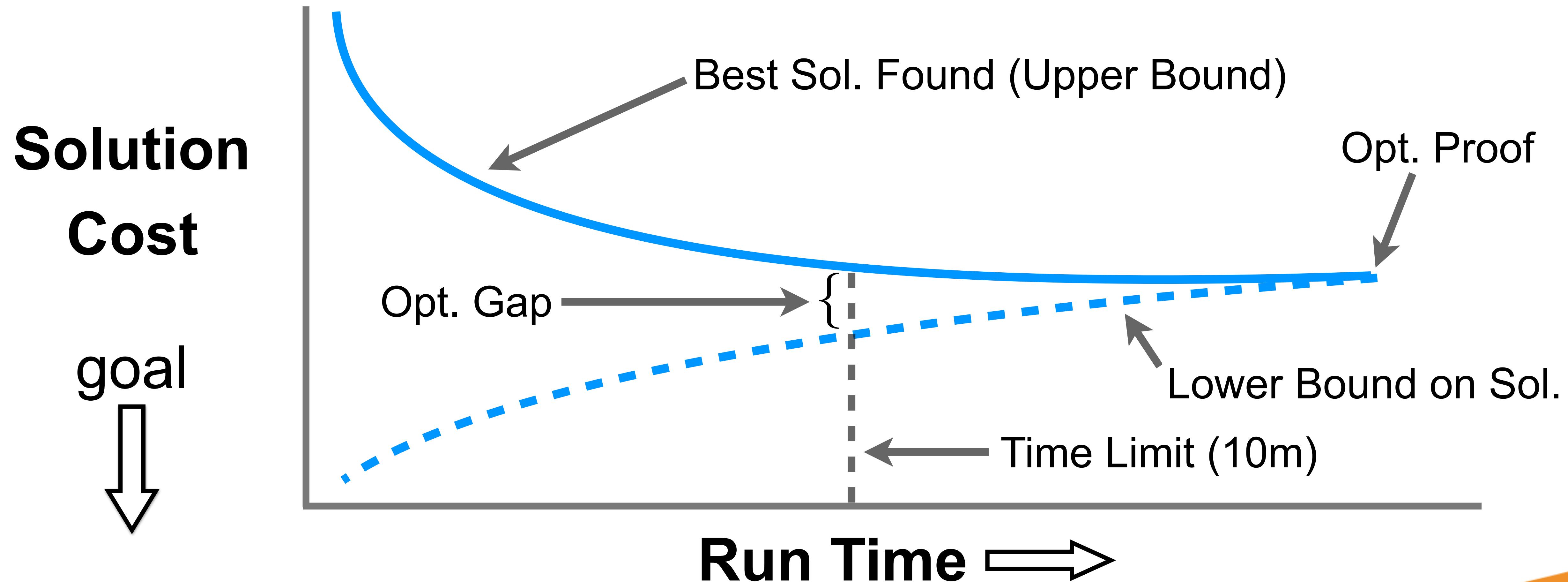


Part 2



C_{12}

Detailed Runtime Studies



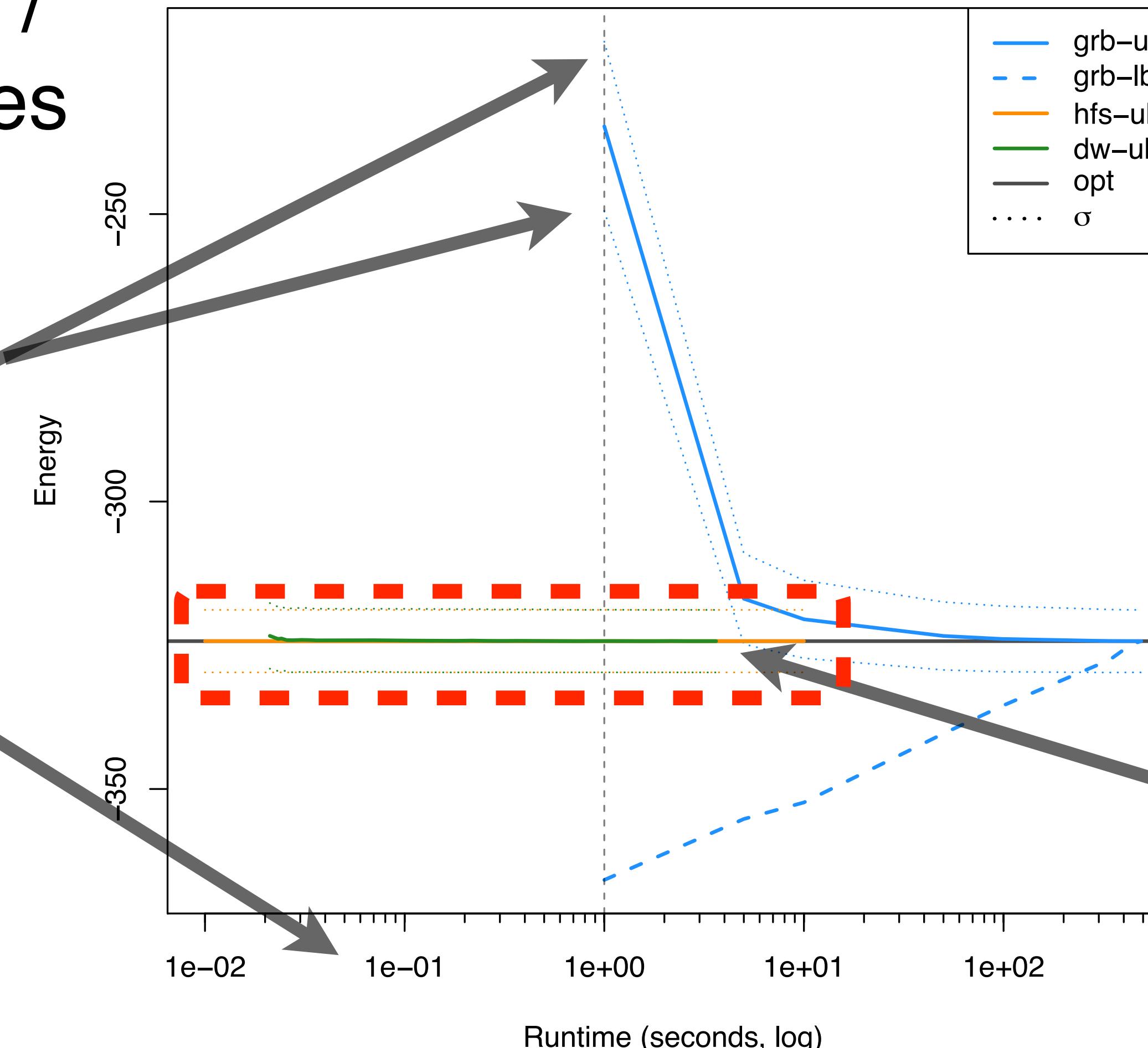
Detailed Runtime Study 1 (RAN-1) C_5

Average Objective /
Energy of 200 Cases

Variance in Cases
not the algorithm

Logarithmic
runtime scale

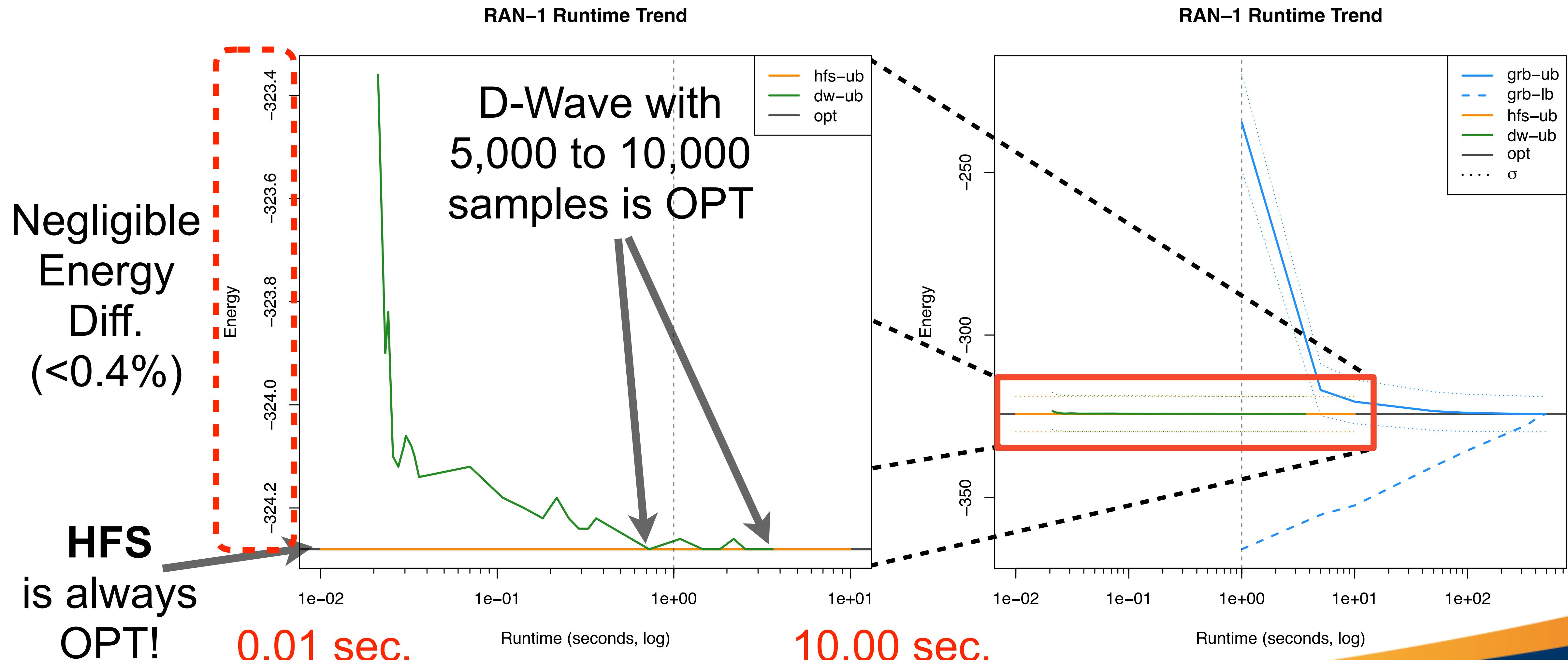
RAN-1 Runtime Trend



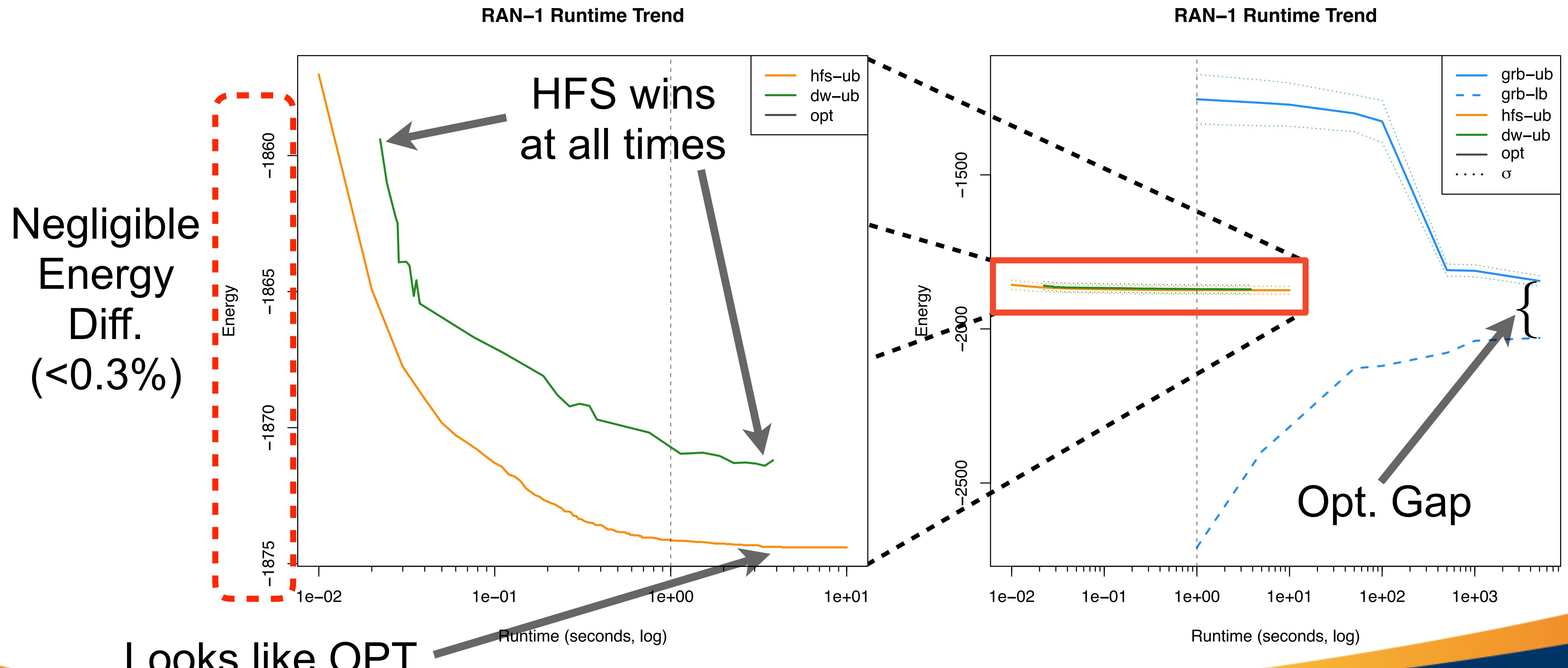
Optimality Proof
for all 200 Cases

HFS and D-Wave
Indistinguishable

Detailed Runtime Study 1 (RAN-1) C_5



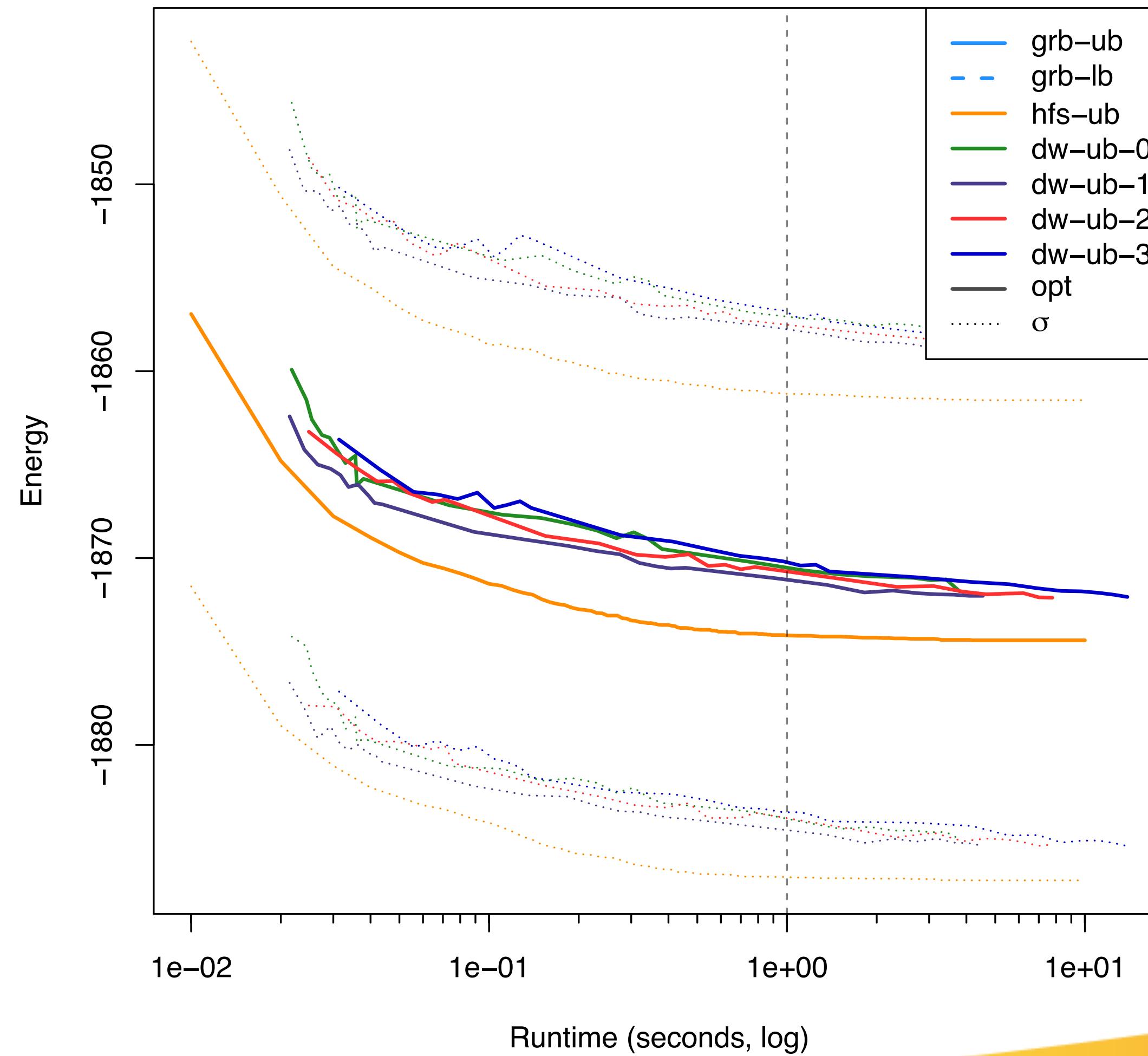
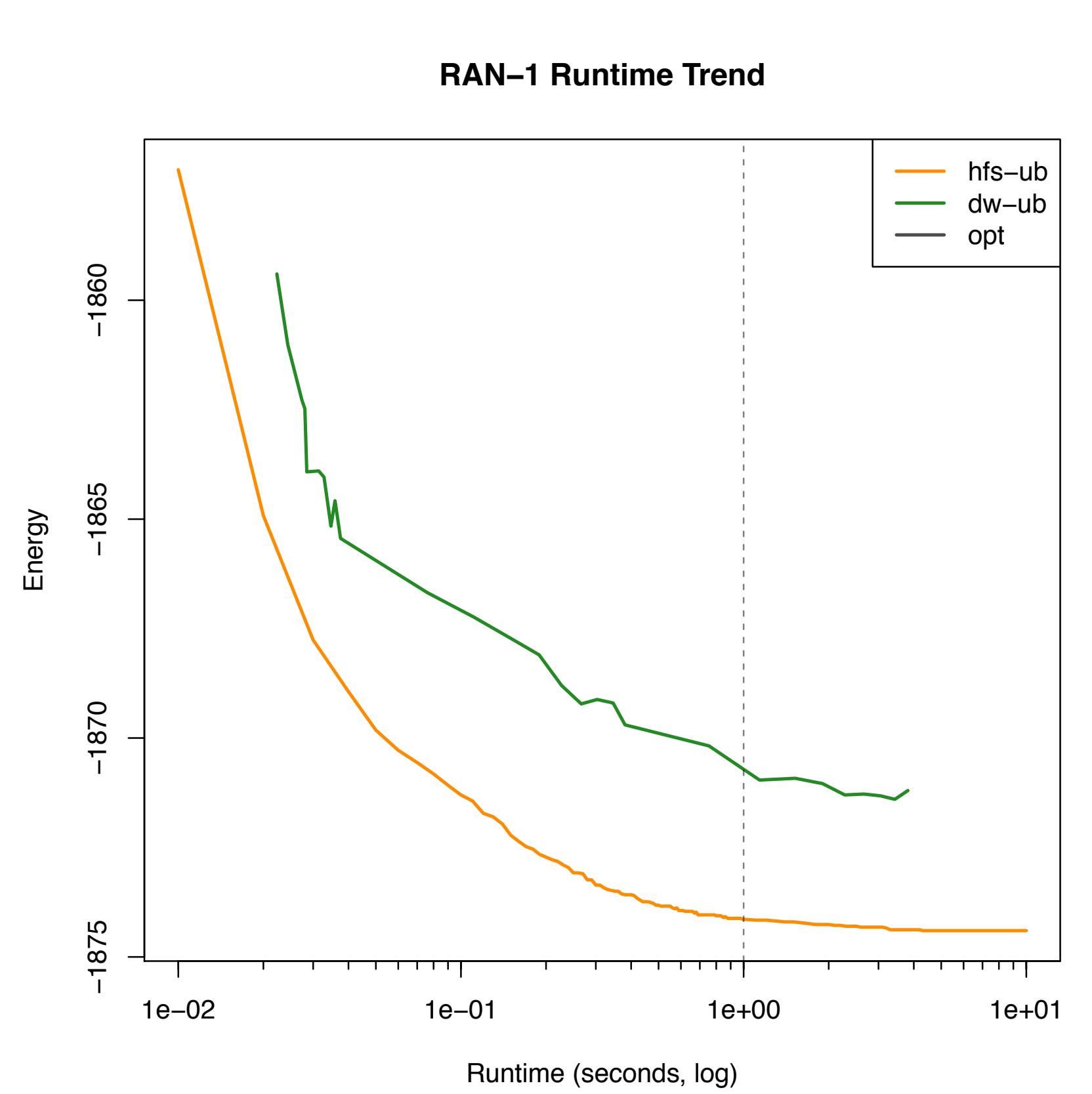
Detailed Runtime Study 2 (RAN-1) C_{12}



Detailed Annealing Time Study

C_{12}

RAN-1 Runtime Trend



- 0 = 5 microsec.
- 1 = 80 microsec.
- 2 = 400 microsec.
- 3 = 1000 microsec.

A Word of Caution about RAN-1

- The relative difference in energy between the **best** HFS solution and the **worst** HFS solution is only **<1%**
 - And a similar property is true for the D-Wave!
- This suggests that the RAN problem has many **nearly-equal local minima**
 - This property is not desirable for **benchmarking heuristics** (e.g. SA, TabuSearch, HFS)
 - Continued work is needed to design generators of **more challenging test cases!**

D-Wave Benchmarking Thoughts

- It seems that all of the popular test cases from the literature are “easy”
- Our D-Wave chip is reliable for well-suited optimization applications (e.g. max cut), but the point where it will overtake state-of-the-art classical methods is not yet clear

Combinatorial Optimization Conclusion

- Combinatorial Optimization is super interesting
- Lot of established algorithms / tools
- QC will have a **BIG** impact (at some point)

Thanks!